Synar Regulation: Sample Design Guidance

May 2003

Synar Regulation:
Sample Design Guidance

May 2003

Contents

Introduction and Requirements

Curtailing easy youth access to tobacco is a crucial component in the primary prevention of tobacco use, and restricting retail sales is an important element of reducing youth access (Clark et al., 2000). Besides helping to reduce current and future health problems among adolescents, compliance with the law is consistent with the public's support of measures to prevent the use of tobacco by young people and, specifically, efforts to discourage tobacco sales to minors (U.S. Department of Health and Human Services [USDHHS], 1994). Changing many facets of the social environment to reduce the broad cultural acceptability of tobacco use is a major recommendation of a recent Surgeon General's report on tobacco (USDHHS, 2000). The report concludes that comprehensive approaches combining community interventions, mass media campaigns, and program policy and regulation are most effective in changing social norms and reducing tobacco use. All of these aspects contribute to creating a healthy environment in which to raise our youth.

The Synar Regulation

In July 1992, Congress enacted the Alcohol, Drug Abuse, and Mental Health Administration Reorganization Act (P.L. 102-321), which includes an amendment (section 1926) aimed at decreasing youth access to tobacco. This amendment, named for its sponsor, Congressman Mike Synar of Oklahoma, requires States to enact and enforce laws prohibiting the sale or distribution of tobacco products to individuals under the age of 18.

Because it plays a lead Federal role in substance abuse prevention, the Substance Abuse and Mental Health Services Administration (SAMHSA) was charged with implementing the Synar Amendment. In January 1996, SAMHSA issued the Synar Regulation to provide guidance to the States. The Regulation stipulates that to comply with the Synar Amendment, each State must do the following:

- Have in effect a law prohibiting any manufacturer, retailer, or distributor of tobacco products from selling or distributing such products to any individual under age 18.

- Enforce such laws to a degree that can reasonably be expected to reduce the illegal sale of tobacco products to individuals under age 18.

- Develop a strategy and negotiate a timeframe with SAMHSA for achieving a retailer violation rate of 20 percent or less.

- Conduct annual, random, unannounced inspections of over-the-counter tobacco outlets and vending machines to ensure compliance with the law. These inspections are to be conducted in such a way as to provide a valid sample of outlets accessible to youth.

- Submit an annual report that details the actions undertaken by the State to enforce its law and includes information on the overall success the State has achieved during the previous Federal fiscal year (FFY) in reducing tobacco availability to youth, the methods used to identify outlets, its inspection procedures, and its plans for enforcing the law in the next FFY.

Failure to comply with the requirements of the Synar Regulation may cause a State to lose a percentage of its Federal block grant funds for substance abuse prevention and treatment. In addition to enacting and enforcing youth tobacco access laws, States' continued receipt of block grant funding depends on reductions in youth access to tobacco as measured by the annual, random, unannounced inspections (also known as the Synar survey). SAMHSA, through its Center for Substance Abuse Prevention (CSAP), Division of State and Community Systems Development, annually reviews each State's plan for conducting the Synar survey and the results, as well as provides technical assistance to States to help them comply with the Regulation.

Objective of the Synar Survey

The objective of the Synar survey is to determine the retailer violation rate (also called the retailer noncompliance rate) for each State based on random, unannounced inspections of a sample of tobacco outlets accessible to youth. The sample of outlets inspected must be representative of the geographic distribution of tobacco outlets in the State. Results of the State's survey will be used to determine whether the State has met the interim target rate negotiated between the State and SAMHSA. The retailer violation rates obtained over a number of years will be used to assess the State's progress toward achieving the overall Synar goal of a violation rate of 20 percent or less. Consequently, it is very important to maintain survey comparability across years. However, if the State or SAMHSA identifies any significant problems or errors, they should be corrected even if doing so may interrupt across-years comparability.

Purpose of This Document

To obtain a valid retailer violation rate, each State must design a sampling methodology that conforms to accepted sampling design practices and fits both the unique circumstances of the State and the requirements of the Synar Regulation. This document is intended to assist States in enhancing their Synar survey methodology and analyzing the results. It presents sampling design requirements and step-by-step guidelines and suggestions for meeting the requirements.

The document is written for State Synar contact persons responsible for implementing the survey and reporting the results. Detailed technical sections present explanations, formulas, and examples that are intended to guide the State's sample design expert and statistician in enhancing the State's sampling methodology.

Since 1996, when SAMHSA published the first *Synar Regulation: Sample Design Guidance*, States have become more experienced in their sampling plans and inspection protocols, and the nature of their questions, issues, and problems has evolved. This updated publication provides more details in some areas than the previous version. It also intends to present material in an easier-to-understand format and includes examples of procedures developed by different States to comply with the Regulation.

Survey sampling procedures are based on accepted statistical theory, but practical considerations and context are also important determinants of sample design. Most successful sampling finds a balance between theoretical purity (the ideal) and practical realities. For sampling theory to be

practically applied, it must be tailored to the particular issue being studied and to practical circumstances. In the case of the Synar Regulation surveys, for example, the degree to which tobacco outlets are geographically distributed will, to some extent, determine the type of design that is most appropriate.

It is impossible to anticipate all the difficulties that may arise while conducting these surveys. Each State is different in terms of its population, geography, and available resources. However, we hope that the guidelines presented and discussed in this document will help State authorities apply sound statistical sampling techniques to each problem they face and ensure compliance with the Synar Regulation. The variety of examples included should help States deal with a range of conditions.

Organization of This Document

This document is divided into six steps that will guide users in developing and implementing the Synar survey and in analyzing the results. SAMHSA's requirements and guidelines for how to comply with the Synar Regulation are discussed within each step. The six steps are presented in chronological order and may be used in a step-by-step fashion to design a new sampling plan. Alternatively, users may refer to specific steps and the corresponding guidelines, formulas, and reporting requirements to enhance the existing sampling methodology and reporting procedures. Appendix A presents two examples of sampling methodologies. Appendix B contains guidance on obtaining technical assistance from SAMHSA. Appendix C contains a glossary, appendix D contains a sample checklist for use by State program directors in designing and implementing the Synar survey, and appendix E contains a bibliography.

SAMHSA Requirements

As stated in the Synar Regulation, "The State shall conduct annual, random, unannounced inspections of both over-the-counter and vending machine outlets. The random inspections shall cover a range of outlets (not preselected on the basis of prior violations) to measure overall levels of compliance as well as to identify violations. Random, unannounced inspections shall be conducted annually to ensure compliance with the law and shall be conducted in such a way as to provide a probability sample of outlets. The sample must reflect the distribution of the population under age 18 throughout the State and the distribution of the outlets throughout the State accessible to youth" [section 96.130(d)(1) and (2)].

Although the sample "must reflect the distribution of the population under age 18 throughout the State," experience gained during the first years of implementation showed that a sample that represents the distribution of outlets throughout the State also reflected the distribution of youth and, therefore, is an acceptable way to obtain a valid statewide measure of retailer violations.

Because it is often not feasible to inspect every tobacco outlet in a State for logistical or financial reasons, a random sample of all outlets must be selected for inspection. An acceptable random-sampling plan conforms to accepted sampling design methodology and considers the geographic distribution of tobacco outlets in the State, the costs associated with conducting the survey, and the divisions already in place in the State (e.g., health districts).

Despite the population, geographic, and resource differences among States, SAMHSA has requirements that each State must follow. These requirements are listed below and are also explained in more detail in this document. SAMHSA requires the State to:

1. Request approval from SAMHSA in writing for any changes in sampling methodology prior to implementation of the Synar survey.

2. Develop a sampling frame that includes both over-the-counter and vending machine locations accessible to youth.

3. Develop a sampling frame that includes, at a minimum, 80 percent of the tobacco outlets in the State.

4. Select a sample of outlets to inspect that is representative of the geographic distribution of all tobacco outlets accessible to youth in the State.

5. Design a sampling methodology and implementation plan that are based on sound survey sampling methodology.

6. Estimate the original sample size before implementing the Synar survey.

7. Determine a method of selecting additional outlets to inspect should it not be possible to reach the required minimum number of completed inspections due to sample attrition.

8. Obtain a completion rate of 90 percent or better.

9. Record the actual steps of the survey process in the field, and keep records of all sources of sample attrition in the field.

10. Incorporate the complexity of the sample design as a factor when analyzing the survey results.

11. Weight the results of the Synar survey to account for unequal probabilities of selection, differences in percentages of eligible outlets between strata or clusters, and other deviations from the intended design.

12. Meet Synar Regulation reporting requirements when completing the Annual Synar Report. According to the Regulation, States must provide the following information related to their sampling methodology:

    C   "A detailed description regarding the overall success the State has achieved during the previous fiscal year in reducing the availability of tobacco products to individuals under the age of 18, including the results of the unannounced inspections . . . for which the results of over-the-counter and vending machine outlet inspections shall be reported separately; and

    C   A detailed description of how the unannounced inspections were conducted and the methods used to identify outlets."

Changes in Sample Design

Requirement

- Request approval from SAMHSA in writing for any changes in sampling methodology prior to implementation of the Synar survey.

Overview

A change in sampling methodology is any deviation from the previous year's Synar survey methodology that SAMHSA has not yet approved. This includes, for example,

- Changing from a list frame to an area frame.
- Changing from a stratified to a stratified-cluster sample.
- Changing the age or gender composition of youth inspectors.

Changes in the State's inspection protocol must be approved by SAMHSA in writing before the Synar survey is conducted.

Guidelines for Meeting the Requirement

Guideline 1: Obtain written approval from SAMHSA.

Before States make any changes in sampling methodology, they must obtain written approval from SAMHSA by doing the following:

- Send a written request to the State Project Officer for Synar at SAMHSA/CSAP.

- Describe in detail how the change(s) would be implemented.

- Include the rationale for the change(s) and the possible implications of the change(s).

- Pretest the proposed change(s), if possible, and describe the methodology and results.

- Mention planned changes for the next year in section II of the Annual Synar Report. Also send a written request to and receive written approval from SAMHSA.

Guideline 2: Note unexpected changes.

Any unexpected changes in sampling methodology that occurred in the field should be noted and fully explained in the responses to questions 6, 7, and 8 in the Annual Synar Report. For example, the report must provide explanations for changes in the original sample size (of 20 percent or more) and completion rates (below 90 percent).

Also note and explain unexpected changes in the inspection protocol.

Six Steps for Developing and Implementing the Synar Survey and Analyzing the Results

Step 1: Develop a sampling frame and assess its quality.

Requirements

- Develop a sampling frame that includes both over-the-counter and vending machine locations accessible to youth.

- Develop a sampling frame that includes, at a minimum, 80 percent of the tobacco outlets in the State.

Overview

The Synar Regulation requires each State to develop an adequate sampling frame from which a random sample of tobacco outlets accessible to youth can be selected. There are basically two types of sampling frame: the list frame and the area frame. The list frame is a list of printed or electronic records of tobacco outlets. The list must have up-to-date information on the location of the tobacco outlets in the population. When such a list frame is not available, a sample of geographic areas can be randomly selected so that tobacco outlets within the selected areas are inspected. The list of these geographic areas is called the area frame. Sometimes a list frame is readily available but too incomplete to use it alone. In this situation, a list-assisted area frame can be used to complement the shortcomings of either frame.

Not all tobacco outlets are eligible for the Synar survey. An over-the-counter or vending machine outlet is eligible for the Synar survey if it sells tobacco products and is accessible to youth under age 18. Ineligible tobacco outlets because of access restriction are bars, taverns, or other adult-only clubs that have an enforced minimum age restriction for entry of 18 or older.

The terms "overcoverage" and "undercoverage" are often used to describe the quality of the sampling frame. Overcoverage refers to the inclusion of outlets on the sampling frame that are not eligible for the Synar survey. For example, the frame may include out-of-business outlets, tobacco outlets inaccessible to youth under age 18, or some duplicates. Undercoverage refers to the omission of eligible tobacco outlets from the sampling frame.

Overcoverage is less problematic than undercoverage in the sense that it can be easily assessed during the survey field operation and corrected at the stage of estimating the violation rate. However, overcoverage increases the survey costs, and ineligible outlets that can be identified should be cleaned off a list frame before using it. Overcoverage usually does not exist for an area frame.

Undercoverage not only is difficult and costly to assess but also is difficult or impossible to correct, because good information on undercoverage is not usually available for the Synar survey sample. Therefore, there should be some assurance that the extent of undercoverage is not severe. The undercoverage problem can exist in both the list frame and the area frame if some

---

areas are not covered for some reason (e.g., remote and sparsely populated areas) or field procedures are not faithfully followed.

States are required to assess the quality of their sampling frame to ensure that it has an adequate coverage. States are expected to develop a sampling frame that includes at least 80 percent of the tobacco outlets accessible to youth in the State. This means that the undercoverage rate should not exceed 20 percent. However, SAMHSA recommends a minimum of 90-percent coverage. It should be noted that this coverage requirement is based on the assumption that the sample design itself does not create any further undercoverage. In some cases, the survey design fails to cover some portion of the sampling frame for some reason. This means that some units in the sampling frame have no chance of being selected (i.e., zero selection probability). This is called survey design undercoverage, which adds the undercoverage caused by the incomplete sampling frame. The survey design undercoverage can be more serious than sampling frame undercoverage.

When a new sampling frame is developed, considerable effort is spent in the first year on identifying and assessing the quality of the sources that make up the sampling frame. In the subsequent years, States are expected to have procedures in place to verify that the sampling frame is both accurate and complete (i.e., small overcoverage and undercoverage). The following sections discuss sources that can be used to develop a sampling frame and ways that States can assess its quality. States that have trouble identifying sources and assessing quality may request technical assistance from SAMHSA (see appendix B for guidance on requesting assistance and the types available).

Requirement

- Develop a sampling frame that includes both over-the-counter and vending machine locations accessible to youth.

Guidelines for Meeting the Requirement

Guideline 1: Develop a sampling frame.

List Frame
Several sources may be used to identify tobacco outlets, both over-the-counter and vending machine locations. A list frame may be derived from preexisting lists, which include license lists and business lists. The single best source for developing a list frame is a tobacco license list.

License List
A license list is a State or local listing of licensed retailers of tobacco or alcohol. License lists may be obtained from the State or local licensing agency, health department, or revenue and tax department. In States where licenses are issued by local municipalities, rather than by a central State agency, a list frame can be formed by combining all the local license lists into a centralized database.

Two types of license lists are tobacco license lists and alcohol or food license lists. A list of all retailers licensed to sell tobacco products is frequently the most complete and up-to-date listing of tobacco outlets in the State. If State law requires licensing for tobacco outlets, using this type of list should be the State's first choice because it is usually the most efficient approach.

A list of licensed alcohol or food retailers may include only some tobacco outlets and cannot be used alone. This list should be used solely to supplement other lists. When using a list of licensed alcohol or food retailers to develop their list frame, States must identify the outlets that sell tobacco and are youth accessible.

Business List
If a tobacco license list is not available, a business list may be used to develop or supplement the list frame. Instead of beginning with a list of known tobacco outlets, as is possible with a license list, a business list usually contains diverse retail outlets that are not tobacco outlets. The most common way for States to obtain a business list is to purchase one or more from a specialized company that routinely verifies and updates its lists and can sort them by type of business. This type of list is referred to as a "commercial business list." Another way to obtain a business list is through the State or local chamber of commerce or government agency that maintains business lists. Regardless of the source of the business list, it is important for States to conduct a field survey to assess coverage of the tobacco outlet population.

The following are examples of sources of commercial business lists:[1]

C   American List Counsel (www.amlist.com)
C   Claritas, Inc. (www.claritas.com)
C   Dirmark USA, Inc. (www.dirmark.com)
C   Dun & Bradstreet, Inc. (www.dnb.com)
C   infoUSA, Inc. (www.infousa.com)
C   Survey Sampling, Inc. (www.ssisamples.com)

A key feature of a commercial business list is that each business on the list is assigned a U.S. Standard Industry Classification (SIC) code or North American Industry Classification System (NAICS) code. The code is a numbering system that identifies the type of business. When developing a list frame, the State need not purchase the entire list but only the list of the types of businesses that may sell tobacco products. The following are examples of SIC codes of establishments that may sell tobacco products:

5399	General Merchandise	5999	Miscellaneous Retail Stores,
5411	Food Stores		not classified elsewhere
5541	Gasoline Service Stations	7011	Hotels/Motels
5812	Eating Places	7933	Bowling Centers
5813	Drinking Places	7948	Racing Tracks

---

[1]Reference to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its approval, endorsement, or recommendation by SAMHSA/CSAP.

5912	Drug Stores	7992	Public Golf Courses
5921	Liquor Stores	7996	Amusement Parks
5962	Vending Machines	7999	Amusement and Recreation
5993	Tobacco Stores/Stands		Services, not classified elsewhere
5994	Newsstands		

Note that the above list of SIC's are not exhaustive. Because of this and other quality reasons, a commercial business list is usually inferior to a tobacco license list.

The main governmental statistical agencies have adopted new NAICS codes, and companies that produce the commercial business lists have begun to provide them along with SIC codes. The NAICS codes will be phased into commercial lists and completely replace the SIC codes. For more information about NAICS codes and how they differ from SIC codes, see the U.S. Census Bureau Web site at www.census.gov/naics.

How To Develop a List Frame
• Create a frame file that either represents a single, reliable list source (e.g., tobacco license list) or that combines different sources of the frame if multiple sources are used. Caution: Ensure that every location of chain stores is included.

• Clean the frame file by eliminating duplicates, unusable records (e.g., key information is missing), and out-of-business outlets to the extent possible (e.g., through telephone checks).

• Verify the frame information (address, eligibility, additional locations of chain stores) by telephone contact with outlets on the frame or by other means. Note that every effort must be made not to alert merchants about upcoming inspections, otherwise the State may bias the Synar survey result.

• Assess the quality of the frame (i.e., measure the coverage and accuracy rates).

Area Frame
When an adequate license or business list is not available or an available list frame has poor coverage (i.e., less than 80 percent coverage), the State may consider using an area sampling frame. In this approach, geographic areas in the State are first randomly selected from a list of areas that are mutually exclusive and that cover the entire State. Selection of areas is followed by enumeration, or listing, of all tobacco outlets (over-the-counter and vending machine locations accessible to youth) in each area selected. This listing is accomplished by a visual inspection of all businesses in the selected areas. All of the listed outlets may be taken into the sample for inspection, or a subsample from the list can be drawn. If an area frame is used, the sample design must be a cluster sample design (discussed later in step 2).

Because a complete enumeration of tobacco outlets for all selected areas must be conducted to use an area frame, the survey costs can be substantially higher than the costs associated with using a list frame. Therefore, a list frame is preferable if it is available and has an adequate

coverage. However, States using a list frame should provide empirical evidence for its adequate accuracy and completeness (discussed under the next requirement).

List-Assisted Area Frame
This type of frame can be used to complement the shortcomings of the list frame and the area frame. For example, when a list frame is available but has poor coverage, an area frame can be used along with the available list frame. A survey designed using an area frame is more costly than a survey using a list frame because the eligible tobacco outlets in the sampled areas have to be enumerated in the field. With aid from the existing list frame, this enumeration cost can be reduced.

Guideline 2: Include locations that are accessible to youth.

All tobacco outlets, except bars and adults-only clubs with an enforced minimum age restriction for entry of 18 or older, are considered accessible and therefore eligible and should be included in the sampling frame. SAMHSA recognizes that some eligible outlets may be impractical to inspect, such as those that are considered unsafe for youth to enter. However, it is required that such uninspected outlets must be counted as eligible and included in the count of noncompleted cases.

It is up to the State to identify situations where, in spite of laws restricting youth access, youth under 18 may have access. An example of this is in a bar–restaurant combination where the waitress or waiter may obtain tobacco products from the bar area for a patron of the restaurant. Paramount to developing an adequate sampling frame is an awareness of these situations and the inclusion of such outlets in the frame.

Requirement

- Develop a sampling frame that includes, at a minimum, 80 percent of the tobacco outlets in the State.

Guidelines for Meeting the Requirement

Guideline 1: Update information annually.

The information used in a list frame should be up to date. The sources of the sampling frame (license or business lists) should be kept current by frequent updates and should be updated at least annually.

It is also important that the Synar survey take place soon after the sampling frame is updated. For example, if a list used as a source for a sampling frame is updated annually in August but the Synar inspection sample is drawn in July, the sample is based on a sampling frame that is nearly a year old and may be inaccurate and incomplete.

Guideline 2: Establish quality assessment procedures.

A reliable sampling frame is one of the key ingredients for a successful Synar survey. A sample drawn from an error-prone frame will not reflect the target population, and the survey results will be erroneous no matter how well other survey procedures are executed. Therefore, the State should establish procedures that will ensure the quality of the sampling frame.

Quality assessment procedures will measure the coverage deficiencies of overcoverage and undercoverage. Overcoverage occurs when the sampling frame contains ineligible outlets because of inaccurate frame information. Undercoverage refers to the omission of eligible outlets from the frame and can be a result of inaccurate address information in the frame. Examples of inaccurate frame information are the wrong address of a tobacco outlet and a listed outlet does not sell tobacco products or is out of business.

Frame assessment procedures should provide two important frame quality measures: percent accuracy and percent coverage. Percent accuracy is the percentage of outlets on the sampling frame that actually sell tobacco products and have accurate addresses. The percent accuracy should be estimated and reported. Percent coverage is the percentage of all eligible tobacco outlets that are included on the sampling frame.

Assessment of a List Frame

When a new list frame is developed, the State should conduct a field survey to assess the frame's quality with respect to percent accuracy and percent coverage. Because it would be very time consuming and costly to check the whole list, an area sample may be randomly selected and a survey undertaken only for the sample. Percent accuracy and percent coverage can be estimated from this survey, and the results can be generalized to the whole population. The key objective of this field frame assessment survey is to ensure that the frame covers at least 80 percent of the target population of tobacco outlets.

Overcoverage

If a list frame has an overcoverage problem, a Synar sample drawn from the list will inherit the problem and the frame information for some sample outlets will be inaccurate. The inaccurate information will be discovered during the Synar survey field operation. This verified information must be kept to be used for correction of the problem later at the estimation stage of the violation rate. In the sense that the overcoverage problem can be assessed on a routine basis and remedied, it is less troublesome than undercoverage. However, the sample size must be increased to compensate for the loss from inaccurate frame information. The sample size increase will increase the survey costs; therefore, the frame should be cleaned as much as possible before drawing the Synar sample to reduce the overcoverage. It is also important to have a good estimate of the percent inaccuracy of the final frame to calculate a realistic sample size.

Undercoverage

Most list frames suffer from undercoverage, which is more difficult to assess and handle than overcoverage. Therefore, it is important to ensure that the extent of undercoverage is not severe. A field study conducted before the Synar survey can determine the extent of the list's undercoverage; the requirement is less than 20 percent, but 10 percent is recommended.

To assess a list frame, a State should conduct a frame assessment field study. An area sample is most ideal for this type of study. The area sampling units should be small enough for accurate listing of outlets within them but large enough to make the study manageable. The following steps describe how to conduct such a study.

1. Develop a list of geographic areas (e.g., ZIP Codes) that are mutually exclusive and cover the entire State. (Make sure that the list frame contains this geographic information for all records in the frame.)

2. Select a random sample of geographic areas to be surveyed using an appropriate sample design. To make things simple, a simple random sampling design is recommended.

3. Create a subfile and print the names and addresses of all outlets on the list frame that are located in each sample area.

4. Obtain accurate maps that include the boundaries and streets of the sample geographic areas.

5. For each sample area, determine a route for the field worker to follow so that he or she finds all outlets.

6. The field worker should follow the route and make a list of all eligible outlets in each sample area. If there is any question about whether an outlet is eligible, the field worker should enter the outlet and determine whether tobacco products are sold and whether the outlet is accessible to youth under age 18.

7. For each outlet printed in step 3, the field worker should record the following information: whether the outlet is found in the sampled area, whether it is in business, whether it sells tobacco, and whether it is accessible to minors. If the answer to any of these questions is no, mark the outlet as ineligible with the reason.

8. Compare the list obtained from the field study with the records on the frame file that fall in the same area as follows:

   • Count the number of ineligible outlet records on the frame file for all sample areas (denoted by a). The percentage of this count out of the sample size (denoted by n) is an estimate of the percent inaccuracy (100 minus the percent inaccuracy estimate is the percent accuracy estimate). If a sampling method other than simple random sampling of areas is used, weighting would be needed to get an unbiased estimate.

   • Count the number of outlets found from the field study but not included in the frame file for all sample areas (denoted by b). These outlets constitute a sample of eligible outlets that are not covered by the frame.

- Algebraically, the percent accuracy estimate is given by $100 \times (1 - a/n)$, and the percent coverage rate estimate is given by $100 \times \{1 - b/(n - a + b)\}$. Note that $(n - a + b)$ is the number of outlets that the frame should have in the sampled areas if the coverage is perfect. Again, weighting may be necessary for an unbiased estimation.

The State may want to use the information collected during the study to update the sampling frame. If the State hires an outside vendor to conduct the sampling frame assessment field study, the vendor should not be a provider of the sources of the frame, to ensure that an impartial study is carried out.

Since a business list frame can deteriorate quickly, it is necessary to conduct a frame assessment field study periodically, every 2 or 3 years.

Coverage of an Area Frame
Conceptually, there is no overcoverage when an area frame is used. Undercoverage may exist, however, if during the Synar survey field operation not all eligible outlets are enumerated for inspection or subsampling of the sampled area. To minimize the undercoverage problem, the survey field operation should be closely monitored. In some cases, remote and sparsely populated areas are deliberately left out of the area sampling frame, resulting in known undercoverage. In this case, a good proxy measure of undercoverage can be calculated using the data produced by the U.S. Bureau of the Census on human population. For example, if 10 percent of the State's population live in the areas not covered by the area frame, the outright undercoverage can reasonably be presumed to be about 10 percent. Also, the actual but unknown undercoverage rate is likely to be higher than 10 percent. This kind of undercoverage can occur even when a list frame is used. It should be totally avoided or minimized as much as possible.

Difference between Frame Assessment Study and Synar Survey Area Sampling
The enumeration of eligible tobacco outlets in the Synar survey sample areas is the same as the enumeration operation of the frame assessment study explained on pages 18 and 19; however, the purpose is very different. The purpose of the enumeration in the frame assessment study is to check the accuracy of the list frame and to measure its undercoverage, whereas the enumeration of outlets in the Synar survey using an area sample selects a sample of outlets for inspection. Conducting an inspection during enumeration is cost-effective. If all enumerated outlets are not inspected, a random sample of outlets should be selected from each sample area using a prespecified sampling procedure. It is convenient to use systematic sampling in this case, where every *K*-th outlets found in the area is inspected. The sampling interval *K* should be determined by the sampling procedure (see page 25).

Note that, although more costly, separating the listing operation and the sampling operation is a less error-prone approach, whereby a list of tobacco outlets is first prepared for each sample area by the field staff and transmitted to the central office for subsampling. Subsamples of outlets are then transmitted back to the field staff for inspection.

Step 2: Select a random sample of outlets that reflects the geographic distribution of outlets throughout the State accessible to youth.

Requirements

- Select a sample of outlets to inspect that is representative of the geographic distribution of all tobacco outlets accessible to youth in the State.

- Design a sampling methodology and implementation plan that are based on sound survey sampling methodology.

Overview

If feasible, it is ideal to take a census survey that selects all tobacco outlets in the State for inspection. However, this approach is feasible only when the population size is small, and it is typically used by some of the U.S. territories. For States with a large population, this option is not practical and/or not feasible and the States must use a scientifically acceptable sample survey to meet SAMHSA requirements. This section discusses how to design such a sample survey by using either a list sampling frame or an area sampling frame.

A fundamental characteristic of a scientific sample survey is that it uses a chance (probability) mechanism to select a sample, which is why the sample is called a random sample. Random sampling enables the surveyor to make probabilistic statements about the survey results. The simplest type of random sample is called the simple random sample (see page 25), and sampling outlets and estimation of the violation rate are simple when this sample design is used. For various practical reasons, however, the Synar survey uses a sample design that is more complex than simple random sampling.

Because a sample survey takes only a part of the population, the survey results are subject to an error called the sampling error. A sample design A is said to be more efficient than a sample design B with the same sample size if the sampling error for A is smaller than that for B.

When a sample is designed to conduct a sample survey, two important components are considered: the sampling efficiency measured by the sampling error and the survey cost. These survey components are inversely related. Hence, the survey designer strives to minimize the survey cost while meeting the sampling error requirement (often called the precision requirement) or to minimize the sampling error for a given budget. The former is the case for the Synar survey.

Requirement

- Select a sample of outlets to inspect that is representative of the geographic distribution of all tobacco outlets accessible to youth in the State.

Guidelines for Meeting the Requirement

Two commonly used techniques to conduct a sample survey are stratification and clustering. Both techniques concern the grouping of population elements (tobacco outlets for the Synar survey) into nonoverlapping groups, but they have different purposes and uses.

Stratification

Stratification divides all tobacco outlets into mutually exclusive and exhaustive groups or strata. "Mutually exclusive" means that no outlet belongs to more than one stratum, and "exhaustive" means that no outlet is excluded in the stratification. When the tobacco outlets within strata are similar (homogeneous) with respect to the violation rate, stratification enhances the sampling efficiency. For example, if outlets in urban areas have a higher tendency of violation than those in rural areas, stratification by urban and rural areas will reduce the sampling error. However, stratification is also used for other purposes, such as administrative convenience and a need to obtain survey results by stratum. In these cases, stratification may not help reduce the sampling error.

Strata can be regarded as subpopulations (mutually exclusive) for each of which a random sample is selected independently. Hence, each stratum has its own sample that represents the stratum.

Clustering

Clustering is a technique, similar to stratification, that groups units (e.g., tobacco outlets) into mutually exclusive and exhaustive groups or clusters. Using clusters, when formed as geographically compact groups, requires less time and fewer trips to inspect outlets and, thus, the technique reduces field costs, as compared with other methods. Clustering can enhance sampling efficiency as well, if clusters are formed so that each cluster is an exact or close copy of the whole population in terms of the violation rate.

Unlike stratification, which requires outlets to be similar within a stratum to achieve efficiency, clustering requires that outlets within clusters should be dissimilar (heterogeneous) with respect to violation. Yet commonly used approaches to clustering make it difficult to achieve this, because clustering is often used to create compact geographic chunks of outlets and to select sample outlets through them. This method reduces survey costs. However, because outlets within geographic clusters tend to be similar within the cluster, this type of clustering tends to increase rather than decrease the sampling error.

Difference Between Stratification and Clustering

In terms of grouping of outlets, stratification and clustering are similar, but their purposes are very different. The main and most important difference is that clusters are sampling units, but strata are not. If desired, clusters can be further grouped into strata. Both techniques can be used to enhance the sampling efficiency, although that is seldom the case for the Synar survey. To increase the sampling efficiency, strata should be more homogeneous and clusters should be more heterogeneous in terms of the outlets' tendency of violation. For clusters defined by geographic areas, however, outlets within clusters tend to be homogeneous, and clustering in this case has a negative effect on sampling efficiency.

Guideline 1: Decide whether to use clustering or stratification or both.

Clustering using geographic areas (e.g., counties or ZIP Code areas) will most likely increase the sampling error, but it is cost-effective. If the cost benefit obtained by clustering exceeds the higher cost needed to handle the sampling inefficiency caused by clustering in the form of increased sample size, clustering is beneficial. When the State uses a list frame, clustering is an option; however, the State should consider the sampling-error and field-cost implications. When the State uses an area frame, clustering is not an option and must be used because the areas in the frame are geographic clusters.

After making a decision on clustering, the State should decide whether to use stratification. Stratification usually does not reduce the sampling efficiency, unless very different stratum sampling rates are used; it can be used without serious implications for the cost or sampling efficiency.

Stratification can be used with or without clustering. When it is used with clustering, stratification is performed on clusters (i.e., clustering is performed within strata).

Guideline 2: Choose a sample design.

Stratification and clustering, if used, are only part of a sample design; more has to be done to get a final sampling plan.

A sample design without clustering is simpler because outlets are selected directly. The simplest is the simple random sampling (SRS) design, where a sample of outlets is directly selected with equal probability without imposing any structure. If stratification is used, the design becomes a stratified SRS design, where the SRS method is used in each stratum independently. Other sampling methods, such as systematic random sampling (see page 25), can be used to select outlets directly with or without stratification. These designs are simpler but not cost-effective because the sample spreads out the State.

When clustering is used, the sample design, called a cluster design, becomes more complex. The cluster structure is imposed on the outlet population, and clusters are selected before outlets are selected. When stratification is used on top of clustering, the design becomes a stratified cluster design. The clusters are selected independently from each stratum using a sampling method such

---

as SRS or systematic random sampling. Another useful sampling method to select clusters is the probability proportional to size (PPS) sampling method (see page 26; an example is provided in appendix A). As the name suggests, PPS takes different cluster sizes into account in selection probabilities.

To select outlets from the clusters sampled in the first stage, all outlets in the sampled clusters could be taken. This design is called the one-stage cluster design or the stratified one-stage cluster design if stratification is also used. Instead of taking all outlets, sampled clusters can be subsampled using a sampling method such as SRS or systematic random sampling. This sampling is done in two stages—sampling of clusters at the first-stage and then sampling of outlets from the sample clusters at the second stage—and is called the two-stage cluster design.

If clusters are large, such as counties, smaller clusters (e.g., ZIP Code areas) can be formed within them. These smaller clusters are subsequently selected within selected large clusters before taking an outlet sample. This design is called the three-stage cluster design, and the process can go on in a nested fashion. A cluster design with more than one stage of sampling is called the multistage sample design.

In a multistage sample design, the highest level clusters are called the primary sampling units (PSUs), the second highest level clusters are called the secondary sampling units (SSUs), and so on. The population elements (outlets), which are selected at the last stage, are called the ultimate sampling units.

Clustering by geographic areas is beneficial in reducing the survey field cost. When an area frame is used, the sample design automatically becomes a cluster design because frame areas are clusters.

In general, a cluster design has a lower sampling efficiency than the SRS design. It is often of great interest to compare a complex sample design, such as a cluster design, with the SRS design. For this purpose of comparison, the design effect is used.

Originally, the design effect was defined with respect to SRS without replacement (SRSWOR). However, the definition has changed by using SRS with replacement (SRSWR) and this definition is more commonly used. Therefore, the latter definition will be used in this guide. The abbreviation "SRS" is used to denote SRSWOR in this document unless otherwise specified.

Design Effect
The design effect (often denoted as Deff) of a complex sample design is defined as the ratio of the variance of an estimate obtained from the complex design to the variance of the estimate obtained from the SRSWR design with the same sample size. Let $\hat{\theta}$ be the estimate of interest and D be the complex sample design. Then Deff of $\hat{\theta}$ under D is defined by this formula:

$$\text{Deff}(\hat{\theta}, D) = \frac{\text{Variance of } \hat{\theta} \text{ under } D}{\text{Variance of } \hat{\theta} \text{ under SRSWR}}$$

When D is less efficient for the estimator $\hat{\theta}$ than the SRSWR, Deff $(\hat{\theta}, D) > 1$. The higher the Deff is, the less efficient the sample design becomes for estimator $\hat{\theta}$.

The Deff is very useful in determining the sample size, which is discussed in step 3.

Guideline 3: Decide on a random sampling method for each stage of sampling.

For each stage of sampling, a random sampling method should be chosen. For selection of outlets at the last stage, SRS or systematic random sampling can be used. For selection of clusters, the choice is between SRS, systematic random sampling, and PPS sampling.

Simple Random Sampling (SRS)

SRS is the basic sampling technique in which a group (a sample) of sampling units of a fixed size (say, *n*) is selected in such a way that every possible sample of size *n* is given the same chance of selection. If 5 units are selected from 100, there are $_{100}C_5 = 75,287,520$ possible samples. This is a huge number, and each such sample is given the same chance of selection. SRS may seem complicated, but its main feature is that each sampling unit is equally likely to be chosen, and it is an equal probability sampling method.

The SRS method can be implemented in several ways. For example, the names of all sampling units can be placed in a pool, mixed well, and drawn one at a time until the desired sample size is reached. A method using a random number table is described in most sampling books (e.g., Cochran, 1975), and many computer packages provide an SRS procedure.

The SRS method can be used at any stage of sampling. For example, for a two-stage cluster design, clusters could be sampled by SRS and outlets then subsequently sampled again by SRS.

Systematic Random Sampling

This method selects units systematically from a list of sampling units. First, the sampling interval (denoted by *I*) has to be determined by the integer part of the inverse of the sampling fraction. If *n* units are selected from *N* units, *I* is the integer part of *N/n*. Next, the random start between 1 and *I* should be selected. Let the random start be *k*; the systematic sample consists of the *k*th unit, (*k+I*)th unit, (*k+2I*)th unit, and so on. Systematic random sampling, like SRS, is another equal probability sampling method.

Example

If five units are selected systematically from a population of 100 units, the sampling fraction is 5/100 (.05). The sampling interval (*I*) is the inverse of the sampling fraction, so that for a sampling fraction of 5/100 the sampling interval is 20. After a random start (a random number between 1 and 20) is chosen, every 20th element in the population is selected thereafter. If 2 is the random start, the 5 selected units would be the 2nd, 22nd, 42nd, 62nd, and 82nd unit in the list of 100 units.

This example may be used to illustrate the difference between SRS and systematic random sampling. Once the list is sorted in a certain order, there are only 20 possible samples in this example. Each is completely determined by the random start and has an equal chance of selection. Under the SRS method, however, there are 75,287,520 possible samples. Nonetheless, both methods are equal probability sampling, which means that each sampling unit has an equal probability of selection. The systematic sampling method can be made an SRS by ordering the sampling list totally randomly. This can be achieved by assigning a unique random number (generated by a computer) to each sampling unit, sorting the sampling list by the random numbers, and selecting a systematic sample from it. The systematic random sampling method may be used at any stage of sampling.

When outlets are sorted in a particular order by some outlet characteristics, a systematically selected sample from the list is like a sample obtained from a list stratified by those characteristics. The sample is evenly spread over the implicitly embedded strata in the frame. This type of stratification is called implicit stratification, in contrast with explicit stratification, which refers to the stratification discussed earlier. Implicit stratification is used only in the context of systematic sampling. There is no hard stratum boundary or independent sampling within strata as there is in explicit stratification.

Probability Proportional to Size Sampling
In PPS sampling, the probability of a unit being selected is proportional to the size of the sampling unit. The most appropriate measure of size (MOS) for the Synar survey with a cluster design is the number of outlets in a cluster if a list frame is used. When an area frame is used, this MOS is not available and another size measure should be used. The human population count would be a good proxy size measure.

PPS sampling is feasible only when clusters of unequal sizes are sampled. It cannot be used for outlet selection because outlets do not have a meaningful size measure in the context of the Synar survey. There are many ways to implement this method. One of the most commonly used methods is systematic PPS sampling (an example is provided in appendix A), which is an approximate method but easy to implement.

In PPS sampling, large-size clusters whose sizes are equal to or greater than the sampling interval are selected with certainty; that is, with the probability of 1. In this case, special care should be given to those certainty clusters during variance estimation, where they should be treated as strata. Certainty clusters can be identified before sampling. It is much easier to proceed to separate them out of the PPS sampling procedure and treat them as strata; PPS sampling is then applied to the remaining clusters. Those certainty clusters are sometimes called self-representing because they represent only themselves in providing the overall estimate of the violation rate.

With- and Without-Replacement Sampling
Sampling is usually carried out sequentially by selecting one unit at a time (except for systematic random sampling). If subsequent sampling is done after placing each unit previously selected

back into the sampling pool, the technique is called with-replacement sampling. Here, the same unit can be selected more than once. With-replacement sampling is seldom used in practice. It sometimes is assumed, however, even for a without-replacement sample for variance estimation, because the derivation of the variance estimation formula is much easier for with-replacement sampling. However, without-replacement sampling is more efficient than with-replacement sampling.

Finite Population Correction

In the case of SRS, the variance reduction by without-replacement sampling in comparison with with-replacement sampling is given by a factor of 1 minus the sampling fraction (rate), and this factor is called the finite population correction. If the sampling rate is high, the factor becomes small and consequently the sampling variance becomes small. For example, if the sampling rate is 50 percent, the sampling variance is halved by without-replacement sampling, compared to with-replacement sampling.

Sample Design Features

The sample design tree that follows depicts how a sample design is derived.

SAMPLE DESIGN TREE
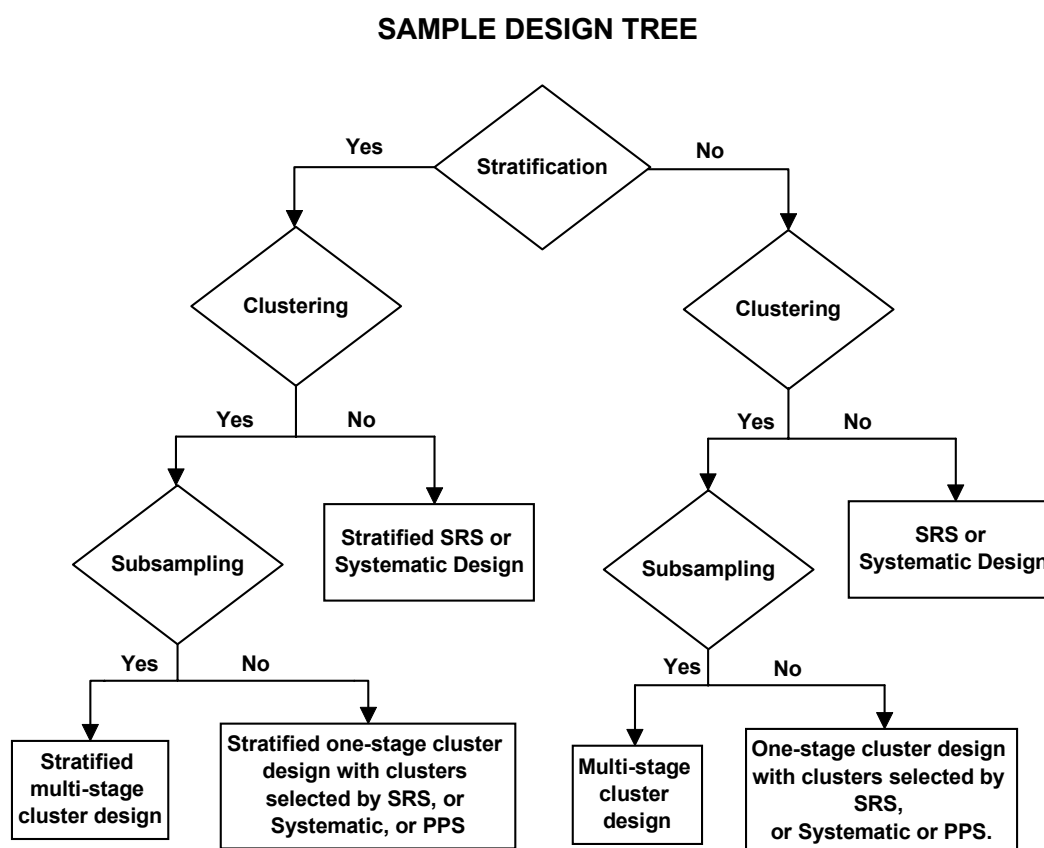
The following design features should be fully explained:

- If stratification is used, explain how strata are formed and the stratum population size. If stratification is done with clusters, give the number of clusters in each stratum.

- When clustering is used, explain how and how many clusters are constructed.

- Explain the sampling method at each stage.

- Explain the population size and the sample size at each stage.

> **Requirement**
> - Design a sampling methodology and implementation plan that are based on sound survey sampling methodology.

Guidelines for Meeting the Requirement

Guideline 1: Ensure that the probability of selection for each outlet is nonzero.

The probability of selection for each unit should be known and above zero at every stage of sample selection. Probabilities of selection should be clearly stated for each sampling unit and stage of sampling. Ultimately, every outlet in the sampling frame should have a nonzero chance of being selected. If some sampling units have zero probability of selection, this constitutes another source of undercoverage, which is different from the frame undercoverage.

In a complex sample design, two main factors contribute to increase the Deff: the clustering effect for cluster sample designs and the effect of unequal sampling weights (i.e., unequal probabilities of selection because the sampling weight is the inverse of the selection probability). Once clusters are formed, one cannot do much about the clustering effect (which can be reduced by using smaller size clusters). However, the weighting factor can be eliminated by using an equal probability sample design of outlets. This is why an equal probability sampling of outlets is more desirable than an unequal probability sampling. For simpler designs without clustering, this sampling can be achieved rather easily. For complex designs with clustering, however, a well thought-out sampling plan at each stage is needed.

The overall probability of selecting a given outlet is the multiple of the sampling probabilities of all stages leading up to the selection of the outlet. Therefore, sampling at various stages must be closely coordinated to achieve the equal probability sampling of outlets. The basic principle is that if the probability of the previous stage sampling is high, the probability of the subsequent stage sampling should be low or vice versa to make the overall product of the probabilities equal.

The following are examples of an equal probability sample of outlets:

1. SRS or systematic random sample of outlets.

2. Stratified SRS or systematic random sample of outlets with an equal stratum sampling rate (i.e., all strata are sampled with a fixed sampling rate).

3. One-stage cluster sample with clusters sampled by equal probability sampling, such as SRS.

4. Two-stage cluster sample, where clusters are selected by equal probability sampling (e.g., SRS) and outlets are sampled from each selected cluster with the same sampling rate. In this scheme, larger clusters have a larger cluster sample size of outlets.

5. Two-stage cluster sample, where clusters are selected by PPS sampling and the same number of outlets is sampled by equal probability sampling (e.g., SRS) from each selected cluster.

6. Stratified one-stage cluster design, where the outlet sampling rate is fixed across the strata and clusters are selected by equal probability sampling (e.g., SRS) within each stratum as in example 3.

7. Stratified two-stage sample, where the same sampling rate (of outlets) is used across the strata and within-stratum cluster and outlet sampling is done as in example 4 or 5.

In a cluster design that has very different cluster sizes, it may be difficult to achieve an equal probability sample of outlets in an efficient manner. Large clusters may have to be broken into smaller clusters, or similar-size clusters may be grouped within strata. Although it may not be possible to obtain an equal probability sample of outlets at the end, it is still desirable to obtain a sample with smaller variation in the selection probabilities (i.e., the sampling weights). It should also be noted that even for the equal probability sample, this feature is often disturbed by differential eligibility rates or completion rates among clusters and across the strata or both (discussed later in step 5). A small variation in the selection probabilities, however, does not cause a substantial increase in the Deff.

Another aspect of a cluster design that should be carefully considered is the size of clusters. In general, the smaller the cluster size, the smaller the clustering effect. However, if the cluster size gets too small, cost saving in the field operation by clustering is reduced as well. Therefore, there should be a good balance between the cluster size and the saving in the field operation.

Guideline 2: Sample eligible vending machines.

Vending machines are often different from the over-the-counter outlets in terms of eligibility and violation rate, and, thus, creating a separate stratum for them would be a good strategy. However, if a State is uncertain whether the number of vending machines accessible to youth is large enough to allow for sampling and for calculating a retailer violation rate separately from the rate for over-the-counter outlets, the State should consult its CSAP Synar Project Officer for guidance.

Step 3: Determine the appropriate sample size for the Synar survey.

Requirements

- Estimate the original sample size before implementing the Synar survey.

- Determine a method of selecting additional outlets to inspect should it not be possible to reach the required minimum number of completed inspections due to sample attrition.

- Obtain a completion rate of 90 percent or better.

Overview

This section discusses the sample size calculations that must be done before implementing the Synar survey, decisions that must be made, and ways to achieve a completion rate of 90 percent or better. The completion rate is the ratio of the number of inspected outlets to the number of eligible outlets in the original sample. A well-executed survey will ensure a high rate of completed inspections.

The minimum number of required inspections is the number of inspections that should be completed to yield reliable survey results. This number is sometimes called the target sample size since it is the targeted number to achieve. The minimum required number of completed inspections is further adjusted to take into account the loss of sample due to the presence of ineligible outlets and a less than 100-percent completion rate. This adjusted sample size is called the *original sample size*, which is the number of outlets a State selects for inspection during the Synar survey. The *final sample size* is the actual number of outlets that were inspected.

Ineligible outlets are dropped from the original sample. An outlet is deemed to be ineligible if the outlet

- Is out of business.

- Does not sell tobacco products.

- Is located in a place that is inaccessible to youth (inaccessible outlets include bars, taverns, and other adults-only clubs with an enforced minimum age restriction for entry of 18 or older).

- Is a private club.

- Is temporarily closed for a period of time (e.g., seasonal closure, closure for renovation).

- Could not be located.

Temporarily closed outlets are sometimes classified as eligible, but the Synar survey is a cross-sectional survey, which means that the survey population is defined for a short time period while the survey is conducted. Therefore, these outlets should be treated as ineligible.

The last category—could not be located—also is not a clear-cut case. The eligibility of an outlet that cannot be located is, in fact, unknown because that outlet may be in business somewhere else. Most of these outlets are believed to be out of business, however, and may be classified as ineligible with a low risk of making a classification error.

Some eligible outlets might not be inspected for the following reasons:

- The outlet was closed at the time of the inspection visit (but is in operation at other times), and the youth inspector was unable to revisit the outlet before the inspection's due date.

- A police officer was in the store at the time of inspection, and the youth inspector was unable to revisit the outlet before the inspection's due date.

- The youth inspector knew the salesperson in the outlet and was unable to revisit the outlet before the inspection's due date.

- The adult supervisor considered the outlet to be unsafe for a youth inspector to inspect.

Requirement

- Estimate the original sample size before implementing the Synar survey.

Guidelines for Meeting the Requirement

Guideline 1: Base the estimate of the original sample size on the results of calculations of the minimum sample size needed to meet SAMHSA's precision requirement plus extra sample needed to account for the expected completion rate and the expected accuracy rate.

Original Sample Size

The original sample size is the number of tobacco outlets the State plans to attempt to inspect during the Synar survey. In developing an adequate sample size, it is almost always necessary to select an original sample larger than the minimum number of required inspections (target sample size) for a variety of factors, including ineligible outlets due to sampling frame inaccuracies and

noncompleted inspections of eligible outlets. Therefore, the original sample size must take into account an expected completion rate and expected eligibility (accuracy) rate that are less than 100 percent. The original sample size will be the minimum number of required inspections plus an additional number of outlets to compensate for the sample loss in the field.

Completion Rate

The completion rate of the Synar survey is the ratio of the number of inspections actually completed (final sample size) to the number of eligible outlets in the original sample that is actually fielded.

$$\text{Completion rate} = \frac{\text{Number of inspected outlets}}{\text{Number of eligible outlets in the original sample}}$$

A high completion rate is beneficial for saving costs and for enhancing the precision of the survey estimate. A low completion may cause a bias problem if not handled properly.

Guideline 2: Determine the original sample size.

The derivation of the original sample size for the Synar survey starts with the calculation of the effective sample size. The effective sample size is the minimum sample size that is needed to meet the precision requirement under the SRS design. However, if the actual sample design is not an SRS design, the effective sample size should be adjusted for the design effect (see page 24 for the definition of the design effect). This adjusted sample size by the design effect is called the target sample size, which should be further adjusted to take the imperfect completion rate and eligibility rate into account to get the original sample size. Note that the sample size does not have to be calculated for a census survey because a census survey takes all the outlets in the population.

The precision requirement for the Synar survey is given in terms of the sampling error. The sampling error of an estimate is measured by the standard error of the estimate, where the standard error of the estimate is the positive square root of the sampling variance of the estimate (variance estimation is discussed in step 5). The sampling variance is inversely related to the sample size under the SRS design, allowing easy calculation of the effective sample size for a given precision requirement.

Effective Sample Size

Let P denote the true violation rate of the State, and let $\hat{P}$ be an estimate of P under the SRS design with an effective sample size n. When the sample size is large (a rule of thumb is $nP(1-P) > 5$) and the sampling fraction (n/N) is small, the sampling theory says that $\hat{P}$ follows approximately a normal distribution with mean P and variance $P(1-P)/n$. If it is required that the standard error (s.e.) be less than or equal to a certain constant (e.g., 0.02), the sample size for the example should be such that

$$\sqrt{\frac{P(1-P)}{n}} \leq 0.02 \;\; \text{or} \;\; n \geq \frac{P(1-P)}{0.02^2} = \frac{P(1-P)}{0.0004}$$

If $P = 0.2$, then n should be at least 400. This is the effective sample size to meet the precision requirement of s.e. ≤ 0.02. In practice, P is unknown, and thus the estimate of the violation rate produced from the prior year's survey should be used instead.
When there is no previous survey estimate, $P = 0.5$ is often used to get a conservative effective sample size. If it is considered too conservative, a good guess value may be used. It would be better to be conservative by picking a value closer to 0.5 rather than closer to 0 or 1. Note that the closer P is to 0.5, the larger the effective sample size is to achieve the same standard error.

Sometimes, $P(1-P) / (n-1)$ is used for the variance of $\hat{P}$, which is theoretically preferable. However, $P(1-P) / n$ is more popular because it is easy to remember, and both are not much different for large n.

If the sampling fraction (often denoted by f; that is, $f = n / N$) is not small (say, > 0.1), the sample obtained from the formula given earlier will be too conservative. In this case, the State may want to incorporate the finite population correction (fpc) given by $(1-f)$ into the variance formula. The variance formula is now the following:

$$(1-f)\frac{P(1-P)}{n} \quad \text{or} \quad \left(\frac{1}{n} - \frac{1}{N}\right)P(1-P)$$

From this, the effective sample size formula is obtained for a given s.e. by

$$n = \frac{1}{\left(\dfrac{(s.e.)^2}{P(1-P)} + \dfrac{1}{N}\right)} \tag{S3.1}$$

If the finite population correction is ignorable, the term $1/N$ in the denominator is ignored and the following simplified formula, which was discussed earlier, is obtained:

$$n = \frac{P(1-P)}{(s.e.)^2} \tag{S3.2}$$

Confidence Intervals
A 95-percent confidence interval for an estimate of the retailer violation rate may be conventionally interpreted as the interval within which the true retailer violation rate would fall 95 percent of the time if the survey is repeated many times. Confidence intervals may be either one sided or two sided, although a two-sided confidence interval is most commonly used. In the case of the Synar survey, where the objective is to determine whether the retailer violation rate is equal to or less than the State target rate, the right-sided confidence interval is more appropriate than the two-sided interval.

The right-sided confidence interval is always bounded by zero on the left. The right-side limit is given by (violation rate estimate) + 1.645 × (s.e. of the estimate). In the following figure, data from a sample Synar report are used to illustrate the right-sided 95-percent confidence interval where the weighted retailer violation rate estimate is 20.97 percent. If the s.e. is 1.55 percent, the right-side confidence limit is 20.97 + 1.645 × 1.55 = 23.52 percent, and the right-sided 95-percent confidence interval is given as [0, 23.52%]. This example may be interpreted as meaning that when the interval is constructed with repeated sampling, it includes the true violation rate 95 percent of the time. This uncertainty results because the violation rate is produced from a sample rather than a census.

One-Sided Confidence Interval
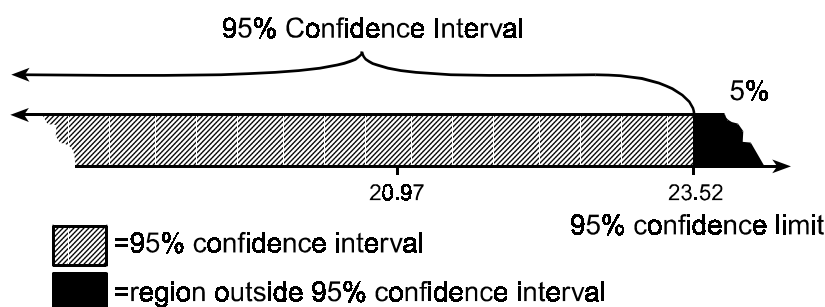
SAMHSA's Precision Requirement for the Synar Survey and the Effective Sample Size
SAMHSA requires that the right-sided 95-percent confidence interval for the estimate of the violation rate has the right-side limit within 0.03 or 3 percentage points from the violation rate estimate.

Using the normal distribution that the violation rate estimate $\hat{P}$ approximately follows, this requirement can be translated into the statement that 1.645 times the s.e. of the estimate be within 0.03. That is, $1.645 \times s.e. \leq 0.03$ or

$$s.e. \leq \frac{0.03}{1.645} = 0.0182$$

(S3.3)

Plugging (S3.3) into (S3.1) or (S3.2), the effective sample size formula is obtained as either

$$n = \frac{1}{\left(\dfrac{(0.0182)^2}{P(1-P)} + \dfrac{1}{N} \right)}$$

(S3.4)

or

$$n = \frac{P(1-P)}{(0.0182)^2} = 3,019\, P(1-P)$$

(S3.5)

Again, *P* should be replaced by the prior year's estimate or a good-guess value if the prior year's estimate is not available.

Example

A State has 34,564 eligible tobacco product outlets. The previous year's violation rate estimate was 24 percent. Because it is expected that the sampling rate is small, the finite population correction (fpc) may be ignored, and the effective sample size is computed by (S3.5) as follows:

$$n = 3{,}019 \times 0.24 \times (1 - 0.24) = 550.66$$

Since *n* should be an integer, round it up to get *n* = 551.

If the fpc is incorporated even though the sampling rate is expected to be small (less than 2 percent), the effective sample size is obtained from (S3.4) as follows:

$$n = \frac{1}{\left(\dfrac{(0.0182)^2}{0.24 \times 0.76} + \dfrac{1}{34{,}563} \right)} = \frac{1}{0.0018160 + 0.0000289} = \frac{1}{0.0018449} = 542.03 \cong 543$$

Note that the effective sample size with the fpc is always smaller than the one without it.

Calculation of the Original Sample Size

To derive the original sample size from the effective sample size, the effective sample size should be adjusted for the design effect, the completion rate, and the eligibility rate.

A complex sample design is usually less efficient, by the factor of the design effect, than the SRS design. Therefore, the sample size for the complex design should be multiplied by the factor of the design effect to meet the same precision requirement.

The design effect is usually unknown, and an estimate or a good-guess value should be used. If the sample design is similar to the prior year's survey design, the prior year's design effect estimate, which was obtained using the definition given on page 24, would be the best. If the sample design is new and the prior year's design effect estimate cannot be used, a good but conservative guess value should be used. For a mildly complex design (e.g., a stratified SRS with mildly different stratum sampling rates), a design effect of 1.3 could be used. For a moderately complex design (e.g., a two-stage cluster design with moderate cluster size (<10) and with moderate variation in the probability of selecting outlets), a design effect of 1.5 would be a reasonable choice. If the design is very complex (e.g., a three-stage cluster design with moderate cluster size (<10) and with quite different outlet selection probabilities), a value of 2 might be an acceptable choice. If the cluster size is large (>10), the clustering effect can be substantial; thus,

a larger design effect should be used. These are rough guidelines, not rules, and should be used with a grain of salt rather than blindly. An experienced survey statistician can get a good-guess value from the past experience of the Synar survey, which would be good enough to use in this context. A useful approximate formula to estimate the design effect in this situation is given in the following:

$$Deff = \{1 + (m-1)\rho\}(1 + CV_w^2)$$

where m is the average cluster size, ρ is the intra-class correlation, and CV_w^2 is the square of the coefficient of variation (or the relative variance) of the sampling weights. For example, if the average cluster size is 10, the inter-class correlation is 0.05, and the relative variance of the sampling weights is 0.3, the approximate design effect is 1.9. If the average cluster size increases to 20 with all others unchanged, then the design effect goes up to 2.5.

From now on, the effective sample size is denoted by n_e. Let the value of the design effect to be used in the sample size adjustment be denoted by d. Then the design effect-adjusted sample size, denoted by n_d, is the following:

$$n_d = dn_e \qquad \text{(S3.6)}$$

For example, if $d = 1.5$, then $n_d = 1.5 n_e$, and so the adjusted sample size is 50 percent larger than n_e. The design effect-adjusted sample size is the required minimum number of completed inspections under the complex design to meet SAMHSA's precision requirement. It was also called the target sample size.

This sample size should be further inflated to compensate for the sample loss because of ineligible outlets in the sample. If the expected eligibility rate (based on the frame assessment study or the prior year's survey results) is denoted by r_l, the eligibility rate adjusted sample size, denoted by n_l, is given by the following:

$$n_l = \frac{n_d}{r_l} \qquad \text{(S3.7)}$$

Finally, the sample size obtained from (S3.7) should be adjusted one more time to get the original sample size (denoted by n_o) by the expected completion rate, for which the prior year's completion rate can be used. Let the expected completion rate be denoted by r_c, and the completion rate adjusted sample size is given by the following:

$$n_o = \frac{n_l}{r_c} \qquad \text{(S3.8)}$$

Combining all the adjustments, a single formula for the original sample size is given by the following:

$$n_o = \left(\frac{d}{r_l r_c}\right) n_e \qquad \text{(S3.9)}$$

It is obvious that to derive the original sample size, all these adjustment factors (the design effect d, the eligibility rate r_l, and the completion rate r_c) are needed. The best estimates or proxies for these factors, of which their true values are unknown, would be the prior year's estimates if the sample design and the sampling frame are similar to the prior year's design and frame. Needless to say, these factors should be accurately estimated each year to be included in the current year's annual report as well as for planning the following year's survey. Nonetheless, they are estimates, and the actual values, especially for the eligibility rate and the completion rate, could be much lower than expected. Therefore, it is prudent to inflate such-derived sample size a little further.

Sample Allocation
After deciding the total sample size, a detailed sampling plan has to be laid out. If the sample design is an SRS, no further detailing is needed; however, this occurrence is rare. In most cases, the sample design is a lot more complex than SRS, and a more detailed sampling plan is needed wherever sampling is conducted. If stratification is used, sampling is carried out independently within each stratum, which necessitates a separate sampling plan for each stratum. This entails the determination of the stratum sample sizes, called sample allocation, and within stratum sample designs.

If an equal probability sample of outlets is desired, the proportional allocation method should be used. This method allocates the total original sample size to the strata proportionally to the stratum population sizes. For example, if a stratum accounts for 20 percent of the whole frame, the stratum gets 20 percent of the total original sample size. Of course, the stratum sample sizes sum to the total sample size. The proportional allocation makes all stratum sampling rates equal.

If the violation rates vary a lot across the strata, Neyman allocation would be more efficient in terms of the sampling error. This method calls for allocating the sample proportionally to the multiple of the stratum population size and the stratum population standard deviation (the positive square root of the stratum population variance). Note that the stratum population standard deviation is given by the square root of the product of the stratum violation rate times the stratum compliance rate (1 – the stratum violation rate), which is unknown, but an estimate from the past survey can be used. Therefore, a stratum with large population size and large standard deviation gets a larger sample size. This allocation is also optimum if the inspection cost per outlet is the same across the strata, in the sense that the method minimizes the sampling variance for a given survey cost. However, an equal probability sample of outlets cannot be obtained from this allocation.

If clustering is used without stratification, how and how many clusters are created and selected should be decided. If outlets are to be subsampled within the selected clusters, within-cluster sample size should be determined. If an equal probability sample of outlets is desired from a cluster design, it can be achieved two ways. One way is to select clusters by PPS sampling and then select an equal number of outlets from each selected cluster (i.e., the within-cluster subsample size is constant and the within-cluster sampling rates vary). In this case, the within-cluster subsample size should be first determined taking the size of the smallest cluster into consideration: If the within-cluster subsample size is too large, it may not be possible to select

that many outlets from small clusters. Once this is done, the number of clusters to be selected can be obtained easily by dividing the total original sample size by the within-cluster subsample size. The clustering effect on the sampling efficiency is reduced as the within-cluster subsample size is smaller. However, a small within-cluster subsample size increases the number of the sampled clusters and in turn the inspection costs. Therefore, some balance between the inspection costs and the sampling efficiency should be considered when determining the within-cluster subsample size. It is always easier to work out the sample allocation when cluster sizes are similar. If they are too variable, it would be difficult to get a desired sample allocation and an equal probability sample of outlets. Breaking large clusters into small ones or combining small clusters into larger ones or both may be needed. An advantage of the method described here is that the resulting outlet sample size is always equal (or close) to the one originally planned. Another important advantage is that the field workload per cluster is equal for all clusters.

The other way of achieving an equal probability sample of outlets is to select clusters by SRS and subsample outlets in the selected clusters with the same subsampling rate in each selected cluster (again by SRS or other equal probability sampling method). In this case, the within-cluster sampling rate should be determined first and then the number of clusters to be selected is given by dividing the original sample size by the average within-cluster subsample size (the within-cluster sampling rate times the average cluster size). This method is easier to implement even when the cluster sizes are very different. However, an important disadvantage is that the resulting outlet sample size cannot be guaranteed to be equal to the originally planned size, but a small deviation of the realized sample size from the one originally planned may not matter much. Another disadvantage is that the field workload is not equal across the clusters.

If stratification is also used with clustering, the discussion above applies stratum by stratum.

Requirement

- Determine a method of selecting additional outlets to inspect should it not be possible to reach the required minimum number of completed inspections due to sample attrition.

Guideline for Meeting the Requirement

Guideline: Select a reserve sample to counter unexpectedly low eligibility and completion rates.

To handle this situation, an extra sample on top of the original sample is selected and put in a reserve (a reserve sample) while the original sample, including the allowance for imperfect eligibility and completion rates, is selected. The reserve sample is used only when needed because field experience indicates a lower than expected eligibility rate or completion rate or both. If the overall eligibility and completion combined rate is about right, even if each is individually different from the expected rate, the reserve sample need not be used.

Twenty percent of the original sample size would be adequate for the size of the reserve sample in most situations. It should be selected from the remaining outlets after the original sample is selected using the same sampling procedure used for the selection of the original sample. Alternatively, first select a sample of increased size that includes both the original and the reserve samples, and then randomly divide the sample into two samples (original and reserve).

When any part of the reserve sample is used, it should be treated as a part of the original sample in every respect. This entails an increase of the sample size and the selection probability in turn. For example, if 10 outlets were selected into the original sample from a cluster with 120 outlets and 2 additional outlets were used from the reserve sample, the original within-cluster selection probability of an outlet was 1/12. Now the sample size of the cluster is increased to 12, and the outlet selection probability is also increased to 1/10. This is a very important point to be noted because States often treat the reserve sample differently, in which case the correct sampling weights needed to calculate the weighted violation rate estimate cannot be obtained.

The treatment of the reserve sample used in the field in this way is quite different from the replacement (or substitution) method, whereby ineligible outlets or noncompletion cases are replaced by outlets from the reserve sample. The former adds a new sample to the original sample, resulting in an increased sample size. The latter simply replaces the unusable original sample outlets by reserve sample outlets, resulting in no change in the sample size. Recent advances in the sampling theory strongly advocate the use of the former method; the substitution method is considered obsolete and should be avoided.

The handling of the reserve sample is sometimes cumbersome. Therefore, some States prefer increasing the original sample size 10 or 20 percent at the beginning. Of course, this strategy is usually more costly but less error-prone and much easier to implement.

Requirement

- Obtain a completion rate of 90 percent or better.

Guidelines for Meeting the Requirement

Guideline 1: Allocate adequate resources.

A high rate of completed inspections is needed to minimize nonsampling error. A large part of survey resources, training, and overall effort should be directed to increasing the completion rate as high as possible. Although many statistical procedures have been developed for proper handling of nonrespondents (noncompleted inspections of outlets), these procedures are remedies, which are usually imperfect, to a problem that is much better solved in the field.

Guideline 2: Do a pretest.

Do a one-time, small pilot test to improve all the survey procedures, which entails executing all the survey procedures with a very small sample. It should be found very useful when a new survey is launched. It is a good idea for the Synar survey as well when a new procedure is implemented. Pretesting enables the survey design staff to see unexpected and unforeseen situations and errors and provide appropriate measures to handle them when they occur in the field. A small investment up front can save a lot of time and money that may be needed to fix the problems that will occur because of poorly designed survey procedures. A well-executed survey will facilitate the goal of having a high rate of completed inspections.

Guideline 3: Closely monitor the survey field operation.

The survey field operation should be closely monitored to detect any problems at the beginning when they occur and to correct them, and also to make sure that the field operation is going smoothly as planned. Special attention should be paid to ensure a 90-percent completion rate at a minimum.

Step 4: Implement and monitor the Synar survey.

<div style="border:1px solid black">

Requirement

- Record the actual steps of the survey process in the field and keep records of all sources of sample attrition in the field.

</div>

Overview

States are required to keep detailed, accurate, well-organized, and preferably computer-based records of all steps in the survey process. Also, States should tally the final dispositions for all outlets assigned to each inspector and report the summary. The final disposition should indicate whether the sample outlet was eligible and the reason for ineligibility if it was ineligible (see page 31 for possible reasons) and whether it was inspected and the reason for noncompletion if it was not inspected (see page 32 for possible reasons of noncompletion).

The detailed information on the disposition of each sample outlet makes it possible to weight the sample so that the events that occurred in the field are properly reflected, and nonsampling error is minimized. The nonsampling error is an error caused by any survey procedures or survey events other than sampling. The sampling error can be reduced by increasing the sample size. However, nonsampling error is not a function of the sample size; it may remain large even when the sample size approaches the population size.

The sources of the nonsampling error include imperfect frame, improper execution of survey protocol, data entry error of the survey results, improper handling of ineligible and noncompleted outlets, and improper calculation of the weights for estimation. As can be seen here, the nonsampling error can creep in at every step of the survey operation. Moreover, it is difficult to quantify. Therefore, to control the nonsampling error, well thought-out quality assurance procedures should be incorporated into the whole survey operation from the frame building to the report writing. The field operation is the most error-prone part of the survey process and so requires good planning and close monitoring.

Guidelines for Meeting the Requirement

Guideline 1: Record all sources of attrition in the field.

States should keep a tally not only of all ineligible outlets that were found and the reasons for ineligibility but also of all eligible outlets that were not inspected and the reasons for noncompleted inspections. Ineligible outlets and noncompleted inspections of eligible outlets should be recorded by cluster for a cluster sample design or by stratum if stratification was used without clustering. The specific reasons why planned inspections were not conducted should be recorded, and the tally should record the outcomes by outlet and by inspector. A sample tally sheet follows.

Sample Tally Sheet

No.	Name of Outlet	Name of Adult Inspector	Youth Inspector			Disposition[1]	Ineligibility[2]/ Noncompletion[3]	
			Name or ID code	Age	Sex		Code	Further Detail

[1]Disposition
  I: Ineligible
  EN: Eligible and noncompleted
  EC: Eligible and compliant
  EV: Eligible and in violation

[2]Ineligibility
  I1:   Out-of-business
  I2:   Does not sell tobacco products
  I3:   Inaccessible by youth
  I4:   Private club
  I5:   Temporary closure
  I6:   Unlocatable
  I7:   Other (explain)

[3]Noncompletion
  N1:   In operation but closed at the time of visit
  N2:   Unsafe to access
  N3:   Presence of police
  N4:   Youth inspector knows the sales person
  N5:   Other (explain)

The sample tally sheet is useful to analyze the survey results by adult and youth inspectors to detect any peculiar patterns. It is often beneficial to examine closely such peculiarities to find causes so that improvement in the inspection protocol can be made. For the reporting purpose, the sample tally sheet should be summarized as shown in the summary sample tally table on the following page.

The table shows that the original sample size is 737 (= 639 + 35 + 63), of which 63 outlets are ineligible and 674 are eligible. The unweighted eligibility (accuracy) rate is then 91.6 percent. Of those 674 eligible outlets, 35 were not inspected and thus, the completion rate is 94.8 percent. These rates along with the summary sample tally should be reported to CSAP/SAMHSA in the Annual Synar Report.

States often make mistakes in the classification of ineligible outlets as non-completed eligible outlets and vice versa. Such miss-classification has a direct effect on the estimation of the retailer violation rate and special attention should be paid to the classification.

Guideline 2: Handle low eligibility and completion rates.

The original sample contains additional outlets to counter the sample loss because of ineligible outlets and noncompleted inspections of eligible outlets. If the actual eligibility and/or completion rates are lower than expected, however, the original sample is not big enough to get the required minimum sample size. Consequently, a reserve sample is selected in advance and used in the field as necessary. However, the release of the reserve sample should be controlled by the central office because it should be used only when necessary. To do this, the progress of the field operation should be frequently and closely monitored. The central office also should provide the field staff with a detailed outlet inspection protocol, which should specify how many

Summary Sample Tally

Result Code	Description	Eligible Complete	Eligible Non-Complete	Ineligible
		Count		
EI	Outlet is eligible and inspected	639		
N1	Outlet is in operation but closed at the time of visit		14	
N2	Outlet is unsafe to access by youth inspector		16	
N3	Police is present in the outlet		2	
N4	Youth inspector knows the sales person		3	
N5	Other (explain)			
I1	Outlet is out of business			20
I2	Outlet does not sell tobacco products			11
I3	Outlet is inaccessible by youth			8
I4	Outlet is a private club			3
I5	Outlet is closed for a period of time			5
I6	Outlet cannot be located			12
I7	Other (explain)			4
Total		639	35	63

inspection attempts the field staff should make before assigning a final disposition of noncompletion to an outlet.

Guideline 3: Supervise field staff.

As with any large-scale and complex operation involving teamwork, it is imperative that rigorous monitoring and quality control be included as an integral and important part of the survey process. The work of all field staff, including adult supervisors and youth inspectors, should be regularly checked for errors. Oversight should not be an afterthought or marginal activity for survey managers and should be included in planning and budgeting.

Guideline 4: Follow the State's approved tobacco outlet inspection protocol.

Each State is responsible for developing a standardized inspection protocol for performing Synar survey inspections. Tobacco outlet inspection protocols include the State's procedures for recruiting, selecting, and training adult supervisors and youth inspectors. The inspection methodology should be clearly articulated and include how the State chooses to conduct its inspections, such as consummated or unconsummated buys, instructions for carrying and showing identification, team composition, and whether the adult supervisor enters the outlet with

the youth inspector. The protocol clarifies data collection procedures and specifies what followup is to be taken after the inspection takes place. Training the inspectors is extremely important to ensure their faithful observation of the protocols.

It is highly recommended that the central staff responsible for planning and managing the Synar survey accompany the inspection teams on a spot basis, especially at the start of the survey, to ensure that the teams understand and follow the procedures correctly.

Although all States have SAMHSA-approved inspection protocols, improvements are always possible. Changes in the protocol require review and written approval by SAMHSA. Adhering to the standardized protocols helps ensure that retailer violation rate estimates are comparable over time (and, ideally, with those of other States). Some important aspects of inspection protocols follow:

- **Youth inspectors.** States should recruit a sufficient number of youth inspectors to allow an approximately equal number of inspections by both age and gender of youth inspectors. Even distribution of outlets to the youth inspectors by age and gender is important because it is generally understood that age and gender have an effect on the buy rate. It is also important that the same distribution is maintained from year to year to make a valid year-over-year comparison.

- **Training.** States should ensure that all members of the inspection team have adequate training.

- **Random assignments.** States should randomly assign youth inspectors to different inspection sites, keeping in mind that the racial and ethnic characteristics of inspectors should reflect site assignments in certain neighborhoods. Randomly assigning minors helps reduce the bias due to differing characteristics of the inspectors. Alternatively, as done in some States, a core group of youth inspectors is recruited and trained to travel across the State to conduct all inspections. These and other characteristics of the inspection protocol should be kept as consistent as possible from year to year to ensure that year-to-year retailer violation results are comparable.

Readers should refer to *Implementing the Synar Regulation: Tobacco Outlet Inspection* for information about the tobacco outlet inspection protocol. This guidance document presents information on methods for performing inspections of retail tobacco outlets as required by law and will assist States in developing a more scientifically credible and reliable protocol. The document is written from the standpoint that the immediate goal of each Synar inspection is to determine whether, under typical circumstances, a tobacco retailer will sell to a minor. The document covers the issues surrounding the use of minors to conduct inspections of tobacco outlets, discusses various data collection approaches, and provides a sample inspection protocol with minors as participants.

Step 5: Analyze the results of the Synar survey.

<div style="border:1px solid black;">

Requirements

- Incorporate the complexity of the sample design as a factor when analyzing the survey results.

- Weight the results of the Synar survey to account for unequal probabilities of selection, differences in percentages of eligible outlets between strata or clusters, and other deviations from the intended design.

</div>

Overview

For reasons of practicality and cost, stratification and clustering are almost always used together or separately when selecting a sample that reflects the geographic distribution of all tobacco outlets accessible to youth in the State. Thus, the sample design is complex, and weighting the results of the surveys will be required in virtually all Synar surveys. It is necessary to account for unequal probabilities of selection and differences in percentages of eligible outlets between strata or clusters and other deviations from the design. It is often beneficial to make some adjustments that bring the sample results more closely in line with known population distributions.

<div style="border:1px solid black;">

Requirement

- Incorporate the complexity of the sample design as a factor when analyzing the survey results.

</div>

Guidelines for Meeting the Requirement

Guideline 1: Use specific formulas.

All formulas used in calculations should be specific to the sampling plan used.

Guideline 2: Calculate a separate retailer violation rate for vending machines when necessary.

If vending machines are legally accessible to youth, a separate retailer violation rate should be calculated for vending machines, along with a standard error. To do this, vending machines should be separated out to form a stratum of vending machines. An independent sample should be selected, and a separate estimate should be calculated.

If a separate violation rate is estimated for vending machines, an estimate of the overall violation rate is obtained by combining it with the estimate for over-the-counter violations by using

appropriate combining factors. The relative over-the-counter and vending machine eligible population sizes are the best choice for the factors. Since the true eligible population sizes are not known, they should be estimated. If the estimated over-the-counter and vending machine eligible population sizes are $\hat{N}_1$ and $\hat{N}_2$ respectively, the relative sizes that can be used as the combining factors are $\hat{N}_1 / (\hat{N}_1 + \hat{N}_2)$ and $\hat{N}_2 / (\hat{N}_1 + \hat{N}_2)$. Note that the combining factors are summed to 1.

When vending machines are not sampled separately but together with the over-the-counter outlets, then a separate estimate for vending machines can still be obtained by using so-called domain estimation technique. However, the sample size for vending machines is often too small to support a separate estimate.

Guideline 3: Calculate the sampling error taking the complex design into account.

Analysis of data collected from Synar-survey inspections would be very straightforward if all Synar surveys were based on simple random samples with no stratification or clustering. Unfortunately, this is almost never the case. For reasons of practicality and cost, stratification or clustering or both are often used. A complex sample design with area-based clusters is less efficient than an SRS design because of the clustering effect. If this design is treated as if it were an SRS sample, the true sampling variance would be underestimated. Thus, results that might seem statistically significant using formulas appropriate for simple random sampling, in fact, may not be. The complexity of the design can be incorporated into the analysis phase by running a specific software program that can handle data collected from a complex design in the calculation of sampling errors. However, depending on each State's individual circumstances, this approach may not be straightforward. States that use a clustered design are encouraged to request technical assistance from their CSAP Synar Project Officer.

Requirement

- Weight the results of the Synar survey to account for unequal probabilities of selection, differences in percentages of eligible outlets between strata or clusters, and other deviations from the intended design.

Guidelines for Meeting the Requirement

To analyze sample survey data, survey weights are needed. The most commonly accepted concept of the survey weight is the representation weight; that is, how many population elements are represented by the sample element including itself. Thus, the sample data can be expanded to the population through the survey weights. Unequal probabilities of selection of outlets for the Synar survey and the sample loss due to ineligibility and noncompleted inspections of some sample outlets can be accounted for by weighting. It should be emphasized that proper weighting is necessary to obtain valid survey results. Detailed discussion on weighting follows.

Guideline 1: Calculate survey weights.

After data are collected in the field, they have to be edited, coded, keyed, cleaned, and analyzed. After cleaned data are produced, survey weights are calculated to account for unequal probabilities of selection and other deviations from the design and to make final adjustments that bring the sample results more closely in line with known population distributions.

Weighting the survey data involves the calculation of the sampling weight for each sampled outlet in the original sample and a series of adjustments to compensate for sample attrition due to ineligible outlets and noncompletion of inspections. Note that a sampling weight is assigned to every sampled outlet in the original sample regardless of whether it is eligible or whether its inspection is completed.

Base Sampling Weight and Adjustments

The base sampling weight of a sampled outlet is the reciprocal of the probability used to select the outlet. If the sample design is simple, such as SRS or stratified SRS (systematic random sampling and stratified systematic random sampling design fall in this category), the base sampling weight calculation is easy, but it must be done stratum by stratum if (explicit) stratification was used. For example, if 20 outlets are sampled from a stratum of 200, each sampled unit has a selection probability equal to 0.1 (= 20/200) and the base sampling weight equal to 1/0.1 = 10, and thus it represents 10 outlets in the stratum. In this case, the outlet selection probability is the same as the within-stratum sampling rate, and the base sampling weight is also given by the inverse of the sampling rate. This is true for SRS but not in general. If another stratum was sampled at twice the rate (i.e., 0.2) the base sampling weight assigned to each of the sampled outlets in the stratum would be equal to 1/0.2 = 5. Note that the sample size for the stratum depends on the stratum population size and the stratum sampling rate. If the stratum population size is 100, the stratum sample size is 20. If the stratum sampling rates were the same at 0.1 for both strata, however, the second stratum sample size would be 10 instead of 20, and the base sampling weight would be the same as for the first stratum. From this example, it is easily seen that if the stratum sampling rates are all the same for all strata under the stratified SRS, the base sampling weight is the same for all sampled outlets.

If the sample design is a complex one with clustering, the base weight calculation must account for each stage of sampling. The overall probability of selecting an outlet is the multiple of sampling probabilities used in all stages. If the sample design is a one-stage cluster design, all outlets within selected clusters are sampled and outlet selection probability within clusters is 1. The overall outlet selection probability is 1 times the cluster selection probability, which is the same as the cluster selection probability. In a two-stage design, where outlets are subsampled within selected clusters, within-cluster outlet selection probabilities are not 1 but less than 1, and the overall outlet selection probability is smaller than the cluster selection probability. For example, if a cluster is selected with a probability of 0.2 and outlets in the cluster are subsampled by SRS with a sampling rate of 0.5, the outlets in the cluster have a selection probability of 0.1 (= 0.2 × 0.5) and their base sampling weight is 10 (= 1/0.1). For a multistage design with more than two stages, this process goes on until all sampling stages are accounted for.

One important characteristic of the base sampling weight is that the total of all base sampling weights is equal (or close) to the frame population size. This total is one simple check of the correctness of the base sampling weight calculation. (When an area sampling is used without any list frame, this is not possible because the population size is not known. However, a population estimate should be available from the previous year's survey results and can be used for the purpose of checking. Of course, in this case an exact equality cannot be expected.)

The base sampling weight is then further adjusted to account for noncompleted inspections. This adjustment entails knowing the tendency of an outlet to have a noncompleted inspection, information that can be gleaned from looking at completion rates across various subclasses. The (unweighted) noncompletion rate is the proportion of all outlets for which inspections were not completed for any reason. The smaller the subclasses are, the better the noncompletion adjustment can be. However, the subclass sizes should not be too small, because small subclass sizes tend to increase the variability of the adjusted weights and increase the sampling variance in turn. Therefore, small subclasses (preferably with similar completion rates) should be collapsed to get a larger subclass. In a cluster sample design, clusters may be used as the adjustment subclasses.

Once the adjustment subclasses are determined, the noncompletion weight adjustment factor should be first computed by subclass. The factor is defined as the ratio of the sum of the base sampling weights of the eligible outlets in the subclass to the sum of the base sampling weights of the eligible outlets with completed inspection in the subclass. Let *w* denote the base sampling weight. Then the noncompletion weight adjustment factor can be written as follows:

$$\text{Noncompletion adjustment factor} = \frac{\text{Sum of } w\text{'s of the eligible outlets in the subclass}}{\text{Sum of } w\text{'s of the inspection completed eligible outlets within the subclass}}$$

The noncompletion adjusted weight is then obtained by multiplying this factor by the base sampling weights of the eligible outlets in the subclass. Note that the adjustment is done within each subclass independently. The adjustment factor is naturally bounded below by 1, but the upper bound is not limited and efforts should be made to control the upper bound. A rule of thumb is that the factor should range between 1 and 2.

Other Beneficial Adjustment
Sometimes, accurate population size information becomes available after the survey was conducted. For example, a survey was conducted using an outdated list frame, but after the survey was conducted, a new updated frame with a high accuracy rate became available. In this case, a further adjustment of the noncompletion adjusted weights can reduce the sampling variance. For this type of adjustment, high-level subclasses such as strata are used. The adjustment factor is calculated by the ratio of the new population size of the subclass to the sum of the noncompletion adjusted weights of all outlets with completed inspection in that subclass. This factor is then multiplied by the noncompletion adjusted weights of all outlets with

completed inspections in that subclass. Note that the adjustment factor in this case can be less than 1, contrary to the other types of adjustment factors.

Final Weights

Note that the weight adjustments discussed above are applied to the final sample of completed inspected outlets. The weight resulting from all these adjustments is called the final weight, and this is the weight used to calculate the weighted violation rate estimate. It is important to understand that the final weights are almost always not equal even for an equal probability sample because of the effect of the adjustment(s). For this reason, weighting must always be done properly by taking into account all stages of sampling as well as noncompletion. Otherwise, some bias will be present in the violation rate estimate. Nonetheless, in the case of an equal probability sample, the final weights should not vary too much because the weighting started with equal base sampling weights, and such a design is desirable for the Synar survey.

Guideline 2: Apply final weights to estimate the retailer violation rate.

Estimate of the Retailer Violation Rate

The final weight must be used to estimate the retailer violation rate. Two totals are first estimated: the total population size and the total number of violating outlets in the population. The total population size is estimated by the sum of the final weights (denoted by $\hat{N}$), and the total number of violating outlets in the population is estimated by the sum of the final weights of the violating outlets in the final sample (denoted by $\hat{Y}$). The weighted estimate of the retailer violation rate, denoted by $\hat{P}$, is given by the following formula:

$$\hat{P} = \frac{\hat{Y}}{\hat{N}} \tag{S5.1}$$

This is a ratio estimate, and although it is slightly biased, it is usually more efficient than other estimators. For example, even when the population size (N) is known (which is rare), the ratio estimate is more efficient than the estimator that uses N instead of $\hat{N}$ in (S5.1).

The estimate given in (S5.1) is the same as the unweighted estimate (the number of violating outlets in the final sample divided by the final sample size) only when the final weights are all the same. However, this almost never happens if the weighting is done properly.

When a stratified design is used, the formula in (S5.1) can be written in a form of linear combination of the stratum weighted violation rate estimates. To write this form algebraically, let $\hat{N}_h$ and $\hat{Y}_h$ denote stratum counterparts of $\hat{N}$ and $\hat{Y}$ (i.e., estimates of the stratum total population size and stratum total violating outlets). First, the combination factor is defined as the relative stratum size as follows:

$$R_h = \frac{\hat{N}_h}{\hat{N}} \qquad h = 1, 2, ..., H$$

Here, the number of strata is assumed to be H. Then the formula in (S5.1) can be written as

$$\hat{P} = \sum_{h=1}^{H} R_h \frac{\hat{Y}_h}{\hat{N}_h} = \sum_{h=1}^{H} R_h \hat{P}_h \tag{S5.2}$$

where $\hat{P}_h$ is the *h*-th stratum weighted violation rate estimate. The $\hat{P}_h$ is the same as the stratum unweighted violation rate estimate only if the stratum final weights are all equal within that stratum. If this is true for all strata, optional Form 2 in the Synar report can be used to calculate the overall State-level weighted violation rate estimate. For other cases, the form could lead to an erroneous estimate and so should be used with caution. The form could still be used for other stratified designs if the stratum weighted violation rate estimates $\hat{P}_h$ is used in place of the stratum unweighted violation rate estimates.

Estimate of the Sampling Variance and Standard Error of the Weighted Violation Rate
The sampling variance of the weighted violation rate estimate $\hat{P}$ is the variance when the sampling process is repeated many times. However, this sampling variance can be estimated based on a single sample.

If the design is SRS and the final sample can be treated as such, the variance estimation formula is very simple:

$$\hat{V}(\hat{P}) = (1-f) \frac{\hat{P}(1-\hat{P})}{n-1} \tag{S5.3}$$

Here, f is the final sampling rate, n is the final sample size, and $\hat{P}$ is the weighted (or unweighted, they are the same in this case) violation rate estimate. If the design is a stratified SRS, this formula is applied stratum by stratum to obtain stratum sampling variance estimates, $\hat{V}(\hat{P}_h)$, and then the total sampling variance is given as follows:

$$\hat{V}(\hat{P}) = \sum_{h=1}^{H} R_h^2 \hat{V}(\hat{P}_h) \tag{S5.4}$$

For a cluster design, the estimate becomes more difficult depending on what sampling method is used to select the primary sampling units. SRS of clusters is easier to handle, but PPS sampling of clusters is more difficult because joint probabilities are needed. To circumvent this difficulty, it is usually assumed that the clusters are selected by with-replacement sampling.

Regardless of the sample design, another complexity comes into play because the estimate is a ratio estimate. This complexity is usually dealt with using Taylor Series expansion or replication methods of variance estimation. (These topics are fully discussed in most sampling texts; see Appendix E: Bibliography.)

For these situations, States should consider using statistical packages that greatly facilitate the calculation of the sampling variance. The most commonly used and well-known systems are WesVar from Westat, Inc., which uses replication methods, and SUDAAN, which is based on the Taylor Series expansion technique. An old version of WesVar can be downloaded from Westat's Web site (www.westat.com/wesvar). SAS (www.sas.com) and Stata (www.stata.com) also have the capacity to calculate sampling variances for complex designs using Taylor Series expansion. To use these packages, however, the survey data set must be appropriately prepared and parameters should be properly specified in the system.

States having difficulties in calculating the sampling variance estimate should consider requesting technical assistance from their Synar Project Officer in identifying a variance estimation methodology that is appropriate to their sampling methodology.

Once the sampling variance is properly estimated, a standard error (s.e.) estimate is obtained by taking the positive square root of the sampling variance estimate. If the estimated s.e. is equal to or less than 0.0182, the State meets SAMHSA's precision requirement.

Confidence Interval

Once the s.e. estimate is obtained, the right-sided 95-percent confidence interval for the weighted violation rate estimate $(\hat{P})$ can be constructed by this formula:

$$\left[0, \left(\hat{P} + 1.645 \times s.e. \right) \right] \tag{S5.5}$$

It is required that the right-side limit of the interval not exceed the noncompliance rate estimate $(\hat{P})$ by more than 3 percentage points.

Meeting the Target Violation Rate

If the weighted violation rate estimate is less than the State's target, SAMHSA recognizes that the State meets the target. Even when the estimate is greater than the target, SAMHSA allows the estimate to go beyond the target by the allowable margin of error given by $(1.645 \times s.e.)$ up to 3 percentage points. Summarizing the requirement as a general rule, SAMHSA recognizes that the State meets the target as long as the State's weighted violation rate estimate does not exceed the critical value determined by the value of the State's target violation rate plus $(1.645 \times s.e.)$ where $(1.645 \times s.e.)$ is less than or equal to 3 percentage points.

For example, if $\hat{P} = 0.223$ and s.e.= 0.0175, then the critical value is $0.2 + 1.645 \times 0.0175 = 0.229$, which is greater than 0.223. Therefore, SAMHSA recognizes that the State still meets the target RVR of 0.2.

Eligibility Rate

The eligibility rate should be calculated and reported as an indicator of the frame accuracy. The sample eligibility rate is simply the ratio of the number of eligible outlets as determined from the field work to the total original sample size including any reserve sample added. This rate applies

to the sample only. However, to use the rate for the next year's survey (i.e., sample size determination) and also as an indicator of the frame accuracy, the weighted rate is more appropriate because it can be generalized to the population. It is defined by the ratio of the total base sampling weights of eligible outlets to the total population size. Note that the base sampling weights instead of the final weights are used for this purpose.

Completion Rate

The completion rate is the ratio of the number of outlets actually inspected to the number of eligible outlets in the sample. This rate refers to the sample and is a good indicator of how well the field work was done.

Rounding Error

All calculations involve some rounding error. During the course of calculation, keep as many significant digits as possible without rounding and apply appropriate rounding to the final result. All rates are required to report at the first decimal point of percent. For example, if the final estimate of the RVR is 16.123567 percent, then report 16.1 percent.

Reporting

It is very important to provide a complete description of the sample design and all deviations that occurred in the field (e.g., eligibility rate, noncompletion rate), in addition to standard tables, analyses, formulae and other products. As a matter of course, it is necessary to include the standard error estimate (correctly calculated) in the report as well as the right-sided 95% confidence interval.

Instructions for completing Form 1, optional Form 2, Form 3, as well as the narrative sections, are included in the *Guide for Completing the Annual Synar Report.*

Step 6: Report the results of the Synar survey to SAMHSA.

Requirement

- Meet Synar Regulation reporting requirements when completing the Annual Synar Report. According to the Regulation, States must provide the following information related to their sampling methodology:

  "A detailed description regarding the overall success the State has achieved during the previous fiscal year in reducing the availability of tobacco products to individuals under the age of 18, including the results of the unannounced inspections . . . for which the results of over-the-counter and vending machine outlet inspections shall be reported separately;

  A detailed description of how the unannounced inspections were conducted and the methods used to identify outlets."

Overview

States are required to report their sampling methodology and Synar survey results in the annual Substance Abuse Prevention and Treatment (SAPT) Block Grant application. It is recommended that questions 6 and 7 of the Annual Synar Report be completed by the State's sample design expert and statistician. However, before sending the Annual Synar Report to SAMHSA, the document should be reviewed by the lead Synar agency in the State to confirm that results are correctly reported in the narrative and adequately substantiated in Forms 1, 2 (if the latter is used), and 3. All formulas and calculations used to estimate the sample size and analyze the results must be included and referenced.

States are encouraged to contact their CSAP Synar Project Officer if questions arise as they complete their application.

Guidelines for Meeting the Requirement

Guideline 1: Describe the sampling methodology used to conduct random, unannounced inspections.

The sampling methodology should be given in detail even if it has not changed from prior years.

Source(s) and Quality of the Sampling Frame
When describing the sampling methodology, it is important to include the following information about the source(s) and quality of the sampling frame:

C   All sources of data from which the sampling frame is built.

C   When the sampling frame was last updated

C   Procedures used to ensure that the addresses on the sampling frame are accurate

C   Criteria used to determine accessibility of outlets to youth

C   Methods used to verify that outlets identified on the sampling frame actually sell tobacco

C   Methods used to locate tobacco outlets that were not on the sampling frame

C   Estimate of percent accuracy: the percentage of the sampling frame that included outlets that actually sell tobacco and had accurate addresses

C   Estimate of percent coverage: the percentage of all State outlets that were actually included on the sampling frame

Random Selection Process
When describing the sampling methodology, it is important to include the following information about the random selection process:

C   State the type of random sample design used to conduct the Synar survey.

•   If stratification was used, give the full definition of the strata.

•   If clustering was used, give the full definition of clusters and the number of clusters for each stratum if applicable.

•   If a multi-stage sample design was used, for each stage describe the sampling units and sampling procedures (simple random sampling, systematic random sampling, or probability proportional to size (PPS) sampling, etc.).

•   If a reserve sample was used, please describe how the sample was selected and how it was implemented.

Original and Final Sample Sizes
When describing the sampling methodology, it is important to include the following information about the original and final sample sizes:

C   How the original sample size and final sample size were determined

C   An explanation of the difference between the original sample size and the final sample size

C   Minimum number of required inspections (the target sample size)

C   Whether the final sample is representative of the geographic distribution of tobacco outlets in the State

Ineligible Outlets and Noncompleted Inspections
When describing the sampling methodology, it is important to include the reasons why inspections of outlets from the original sample were not completed and whether a reserve sample was used. Information reported in the Annual Synar Report should include a complete summary tally of ineligible outlets and noncompleted inspections, specific reasons for ineligibility, and specific reasons for noncompletion of inspections.

Guideline 2: Address changes in sampling methodology.

Report whether the State's sampling methodology changed in the past year. If changes occurred, it is important to include information on when the sampling changes occurred and to provide a timeline that outlines the implementation of these changes. Also indicate the date on which SAMHSA's approval for the change was obtained.

Guideline 3: Report complete results.

Report the complete results of the Synar Regulation survey inspections conducted during the Progress Year for each SAPT Block Grant application. Calculate the unweighted and weighted retailer violation rate estimates, and round to the nearest 10th of a percentage point as the final reported retailer violation rate.

When reporting the results of the Synar survey, States are required to do the following:

- Round the violation rate estimates to the nearest 10th of 1 percent

- Calculate an estimate of the standard error for the weighted retailer violation rate estimate.

- Calculate the right-sided 95-percent confidence interval for the reported retailer violation rate estimate

States may choose to complete optional Form 2 to calculate the weighted retailer violation rate estimate if a stratified design was used. However, within-stratum violation rate estimates (i.e., $\hat{P}_h$ in equation [S5.2] on page 50) needed in the form have to be calculated using properly derived final weights.

States are required to complete Form 3 (Revised Matrix 7a) to show the distribution of outlet inspection results of attempted and successful buys, broken down by both age and gender of the youth inspectors.

Guideline 4: Verify numbers.

Check carefully to ensure that the numbers in Form 3 match the information reported in the narrative and the numbers reported in Form 1 and Form 2 (or other weighting table) if used. The sum of the numbers listed under "attempted" buys in Form 3 should equal the number of inspections performed as requested in column 3 of Form 1 and column 5 of Form 2. The number of "successful" buys in Form 3 should equal the number of violations in column 4 in Form 1 and column 6 in Form 2. When the previous year's report is used to prepare the current year's report, make sure that all updates are complete and coherent.

Guideline 5: Explain discrepancies.

Explain any discrepancies between the data presented in Form 1, Form 2 (or other weighting table) if used, Form 3, and the retailer violation rate reported in the narrative.

Purpose of the *Guide for Completing the Annual Synar Report*

SAMHSA prepared this guidance document to assist States in completing the Annual Synar Report, which asks States to report on their agreement to have in effect and enforce a State law that makes it unlawful to sell or distribute tobacco products to youth under the age of 18. The guide clarifies the reporting requirements, including the level of detail that SAMHSA desires, which will minimize the need for States to provide additional information during the SAPT Block Grant review process.

The guide also is designed to assist new State and territory staff members in completing the Annual Synar Report, serve as a completeness check for experienced staff, provide for more consistency in Synar annual reports across States, and create opportunities for cross-State analysis.

States are encouraged to contact their CSAP Synar Project Officer if questions arise as they complete their SAPT Block Grant application.

Appendixes

Appendix A: Two Examples of Sampling Methodologies

Two hypothetical examples, which are not realistic, are presented here to provide some easy-to-understand explanations on sampling techniques discussed in other parts of this guidance document.

Example A. Stratified Simple Random Sampling (SRS) Design

Consider a hypothetical State with a list frame of 20,000 outlets. The State consists of three counties (A, B, and C), with 8,000, 8,000, and 4,000 outlets, respectively.

Sampling Frame

- List of outlets based on a combination of lists from commercial and State government licensing sources ($N = 20,000$).

- The list provides each outlet's address.

- A frame assessment field study was conducted to measure the eligibility and coverage rates, and estimated rates are 80 and 90 percent, respectively.

Sample Design

- The three counties in the State will serve as the strata.
- The stratum sample sizes are determined proportionally to the stratum population sizes.
- Within strata, outlets are selected by SRS.
- The sample design is equal probability sampling of outlets.

Sampling Size

- Assuming that $P = 0.2$ (violation rate) and ignoring the finite population correction (fpc), the effective sample size needed to meet the target sampling precision is 484 (see equation [S3.5] on page 35).

- Since the sample design is an equal probability sampling design without clustering, the design effect (Deff) is expected to be about 1 and the effective sample size does not have to change to take account of the Deff.

- From a frame assessment field study, the eligibility rate was estimated to be 80 percent, and the eligibility-adjusted sample size is 605 ($\cong 484 / 0.8$).

- Assuming a completion rate of 90 percent, the original sample size that is obtained is 673 ($\cong 605 / 0.9$), which is rounded up to the nearest multiple of 10 to get 680.

- This sample is proportionally allocated to the strata (i.e., by 2:2:1 ratio), resulting in the stratum sample sizes of 272, 272, and 136.

Sample Selection

- An SRS sample of outlets with the allocated sample size is selected from each stratum independently.

Field Procedures

- Provide the addresses of the selected outlets to the inspectors.

- It is important that the inspectors visit these outlets and no others.

- For each outlet, the inspectors must first determine whether the outlet is indeed eligible. If it is not, the inspector records this fact and the reason for ineligibility and proceeds to the next outlet.

- If the outlet is eligible, the inspector performs an inspection and records the result.

- If the inspection of the outlet could not be completed, the inspector records this fact and the reason.

Estimation

- The sample design is suitable to use Form 2 (of the Annual Synar Report) to calculate the weighted violation rate estimate. The survey results are summarized using Form 2 as follows.

TABLE A (Form 2)
Violation Rate Estimation Table for Example A

Strat.	N	n	n1	n2	x	p=x/n2	N'= N(n1/n)	w= N'/Sum(N')	p*w
1	8,000	272	244	244	47	0.1926	7,177	0.4108	0.0791
2	8,000	272	218	196	25	0.1276	6,412	0.3670	0.0468
3	4,000	136	132	131	29	0.2214	3,882	0.2222	0.0492
Total	20,000	680	594	571	101		17,471	1.0000	0.1751

| Unweighted: 0.1769 | Weighted: 0.1751 |

N—original population estimate of outlets in stratum (includes ineligible outlets)
n—original sample size (number of outlets in sample)
n1—number of sample outlets that are found to be "eligible" (i.e., open and selling tobacco) (n1 # n)
n2—number of sample eligible outlets that were inspected (n2 # n1)
x—number of inspected outlets that failed inspection (x # n2)
p—noncompliance rate (p = x/n2)
N'—adjusted population estimate based on number in sample found ineligible (N' = N*n1/n, N' # N)
w—relative stratum size (w = N'/(sum of N' in column 8))

Variance Estimation

- In this case, the SRS variance estimation formula $p(1-p)/(n_2-1)$ can be used for each stratum.

- The overall variance estimate is then obtained by adding $w^2 p(1-p)/(n_2-1)$ over the strata. The estimate is 0.00025033.

- The standard error (s.e.) estimate as the positive square root of the variance estimate is 0.0158.

- Since $1.645 \times 0.0158 = 0.0260$ is less than 0.03, the State meets the SAMHSA precision requirement.

Right-Sided 95-Percent Confidence Interval

- This interval is given by $[0, 0.1751 + 1.645 \times 0.0158] = [0, 0.201]$ or $[0, 20.1\%]$.

Meeting the Target Violation Rate

- Since the weighted violation rate estimate of 17.5 percent is smaller than the target of 20 percent (say), obviously, the State meets the requirement.

- However, if the general rule is applied to see whether the State meets SAMHSA's violation rate requirement, the weighted violation rate estimate is compared with the critical value that is the target plus $(1.645 \times \text{s.e.})$, which is 22.6 percent. Because the estimate is smaller than this critical value, the State is recognized as meeting the target.

Eligibility and Completion Rates

- Unweighted eligibility rate is 87.4 percent.

- A weighted eligibility rate estimate can be calculated in exactly the same way as done for the weighted violation rate estimate except p is computed by n_1/n. The weighted eligibility estimate is 87.8 percent, which is much higher than expected (80 percent was assumed for sample size determination).

- Unweighted completion rate is 96.1 percent.

- Similarly, a weighted completion rate estimate can be calculated by computing p by n_2/n_1. The weighted completion estimate is 96.1 percent, which is much higher than expected (90 percent was assumed for sample size determination).

- These weighted estimates should be used to determine the original sample size for the following year's survey. In this example, weighted and unweighted estimates are very close, and it does not matter which rates are used for the sample size determination. In general, the weighted estimates are preferred.

Example B. Stratified Two-Stage Cluster Sample Design

The population is the same as in example A, but the design is different. Stratified simple random sampling is considered too expensive to use because it entails sending inspectors to individual outlets scattered throughout the State. Therefore, the counties in example B are subdivided into smaller geographic areas (clusters) that are all different sizes in terms of outlet count. Counties A, B, and C have 160, 200, and 100 clusters, respectively, and the average cluster sizes of the counties are 50, 40, and 40 outlets, respectively.

Sampling Frame

- Same as in example A.

Sample Design

- The three counties in the State will serve as the strata.
- The stratum sample sizes are allocated proportionally to the stratum population sizes.
- Within strata, clusters are selected by probability proportional to size (PPS) sampling.
- Outlets are selected by SRS from each sampled cluster.
- An equal probability sample of outlets is desired.

Sample Size

- Assuming $P = 0.2$ and ignoring the fpc, the effective sample size of 484 is obtained (see equation [S3.5] on page 35).

- Because the sample is a two-stage cluster design, the Deff is expected to be larger than 1; assuming the Deff = 1.5, the Deff-adjusted sample size is $726 \left(= 1.5 \times 484 \right)$.

- Assuming an eligibility rate of 80 percent and a completion rate of 90 percent, a sample of $1,009 \left(\cong 726/(0.8 \times 9) \right)$ outlets is needed and is rounded up to 1,010.

- This is proportionally allocated to the strata, and the stratum sample sizes thus allocated are 404, 404, and 202, respectively.

Sample Selection

- Clusters are to be selected by PPS sampling and outlets by SRS. To get an equal probability sample, a fixed number of outlets should be selected within each stratum. It is decided to select 20 outlets from each sampled cluster considering the workload per cluster.

- The number of clusters to be selected is determined by dividing the stratum sample size by 20 (the fixed cluster outlet sample size). Therefore, strata 1 and 2 need 20 clusters each, and stratum 3 needs 10 clusters. (Rounding is needed to get an integer sample size, and so the actual outlet sample size is somewhat smaller than desired in this case.)

- Clusters are sampled by the systematic PPS sampling method as explained below:

    1. Define a cluster's measure of size (MOS) by the number of outlets in each cluster.

    2. Assign cluster ID to each of the clusters, which uniquely identifies the cluster within strata, and to each outlet belonging to the cluster.

    3. Sort the list frame of outlets by cluster ID. In the sorted list, each cluster ID appears exactly MOS times; that is, once for each outlet in the cluster. The sorted list for stratum 1 with 160 clusters and 8,000 outlets is shown in the following table.

Stratum	Cluster ID	MOS (outlets)	Cumulative Count (outlets)	Selection Probability
1	1	55	55	0.1375
1	2	38	93	0.0950
1	3	80	173	0.2000
1	4	60	233	0.1500
1	5	147	380	0.3675
1	6	95	475	0.2375
–	–	–	–	–
–	–	–	–	–
1	160	75	8,000	0.1875

    4. Calculate the sampling interval (denoted by I) by dividing the total stratum population size by the number of clusters to be selected (as $8,000 \div 20 = 400$ for stratum 1).

    5. Select a random number (denoted by R) between 1 and I (the sampling interval). In the example, $I = 400$ and $R = 50$.

    6. Select the cluster where the random start (R) falls. (In the example above, it would be cluster 1 to which the 50th outlet in the sorted list belongs.)

    7. Select again. The second selection is the cluster where the $(R+1)$-th outlet falls, and the k-th selection is the cluster where the $[R+(k-1)I]$-th outlet is contained. In the example, the second selection is the cluster to which the 450-th outlet belongs; that is cluster 6, and so on. When the end of the list is reached, 20 clusters are selected. Note that the probability of selecting a cluster is proportional to the MOS of the cluster.

    8. Determine the selection probability of cluster i, which is given by mt_i / T. Here, m is the size of the cluster sample, t_i is the cluster MOS, and T is the total MOS ($m = 20$, $t_1 = 55$, $t_2 = 38$, $t_3 = 80$, etc., and $T = 8,000$, for stratum 1 in the example). These probabilities are provided in the above table. Note that the probabilities sum to m, the size of the cluster sample (20 clusters in the example).

- After selecting clusters, sample 20 outlets by SRS from each selected cluster independently. The total outlet sample size is 1,000 and not 1,010 as originally planned because of rounding when the cluster sample sizes are computed. To get a sample with the exact size of 1,010, one more outlet may be selected from each of the 10 largest sampled clusters or from each of randomly chosen 10 sampled clusters.

Estimation

- Within-cluster eligibility and completion rates are found to be similar across the clusters within strata, and weighting adjustments are performed within strata. (Otherwise, weighting adjustments should be done within clusters.) This results in equal final weights within strata, and thus Form 2 is still suitable to use to calculate the weighted violation rate estimate. The results of the survey and estimation calculations are shown in Table B below.

TABLE B (Form 2)
Estimation Table for Example B

Strat.	N	n	n1	n2	x	p=x/n2	N'= N(n1/n)	w= N'/Sum(N')	p*w
1	8,000	404	368	356	70	0.1966	7,287.1	0.4191	0.0824
2	8,000	404	320	304	41	0.1349	6,336.6	0.3645	0.0492
3	4,000	202	190	187	43	0.2299	3,762.4	0.2164	0.0498
Total	20,000	1,010	878	847	154		17,386.1	1.0000	0.1813

Unweighted: 0.1818   Weighted: 0.1813

N—original population estimate of outlets in stratum (includes ineligible outlets)
n—original sample size (number of outlets in sample)
n1—number of sample outlets that are found to be "eligible" (i.e., open and selling tobacco) (n1 # n)
n2—number of sample eligible outlets that were inspected (n2 # n1)
x—number of inspected outlets that failed inspection (x # n2)
p—noncompliance rate (p = x/n2)
N'—adjusted population estimate based on number in sample found ineligible (N' = N*n1/n, N' # N)
w—relative stratum size (w = N'/(sum of N' in column 8))

Variance Estimation

- Because the sample design is a cluster design with PPS sampling without replacement, there is no simple formula. Survey sampling software packages, such as Wesvar or SUDAAN, usually assume that clusters are selected by with-replacement sampling and produce a conservative variance estimate for this type of design. This calculated variance estimate is 0.00023156 and the s.e. estimate, the positive square root of the variance estimate, is 0.01522.

- The s.e. estimate is less than the required limit of 0.0182, and the State meets the precision requirement.

- In this case, the variance estimate with SRS assumption is $\hat{P}(1-\hat{P})/n$, where $\hat{P}=0.1813$ is the estimated violation rate, $n=847$ is the final sample size, and the SRS variance estimate is 0.000175242. Using this variance estimate, the Deff can be estimated as $0.00023156/0.000175242=1.32$, which is smaller than assumed for the sample size determination.

Right-Sided 95-Percent Confidence Interval

- This interval is given by $[0,0.1813+1.645\times0.01522]=[0,0.2063]$ or $[0,20.6\%]$.

Meeting the Target Violation Rate

- Since the weighted violation rate estimate of 18.1 percent is smaller than the target of 20 percent (say), obviously, the State meets the requirement.

- If the general rule is applied to see whether the State meets SAMHSA's violation rate requirement, the weighted violation rate estimate is compared with the critical value that is the target plus $(1.645\times s.e.)$, which is 22.5 percent. Because the estimate is smaller than this critical value, the State is recognized as meeting the target.

Eligibility and Completion Rates

- Unweighted eligibility rate is 86.9 percent.

- A weighted eligibility rate estimate can be calculated in exactly the same way as for the weighted violation rate estimate except p is computed by n_1/n. (Note that the meaning of p has changed to mean the stratum sample eligibility rate.) The weighted eligibility estimate is 87.4 percent, which is much higher than expected (80 percent was assumed for sample size determination).

- Unweighted completion rate is 96.5 percent.

- Similarly, a weighted completion rate estimate can be calculated by computing p (stratum sample completion rate) by n_2/n_1. The weighted completion estimate is 96.5 percent, which is much higher than expected (90 percent was assumed for sample size determination).

- These weighted estimates should be used for the future survey. In this example, weighted and unweighted estimates are very close, and it does not matter which rates are used for the sample size determination. In general, the weighted estimates are preferred.

Remarks

- Sorting the list of outlets for sampling of clusters by cluster ID establishes a certain order of clusters in the sorted list. If desired, an implicit stratification can be implemented by defining the cluster ID in such a way to achieve it. For example, if the clusters are ZIP Code areas and the cluster ID is defined by ZIP Code, the geographic stratification is implicitly embedded in a systematic sample of clusters, which is likely to yield a sample of clusters scattered throughout the county stratum because of the implicit stratification.

- If a cluster is of a size greater than the sampling interval (400 for stratum 1), the cluster is selected with certainty because it will be selected no matter what the random start value is. Certainty clusters distort the PPS sampling, and therefore it is easier to handle them by separating them out from the PPS sampling procedure. The certainty clusters are self-representing and are like separate strata within strata, and so may be considered as substrata.

- If the list is believed to be outdated and misses many eligible outlets, the State may want to use list-assisted area sampling. In this case, the field operation will be quite different (see steps 1 and 2 for more information).

Appendix B: Technical Assistance Available to States To Enhance Compliance With the Synar Regulation

Any State may request technical assistance (TA) from the State Prevention Advancement and Support (SPAS) project to meet the requirements of the Synar Regulation. The available areas of expertise, many of which are addressed earlier in this document, may include any of the following:

C   Sampling design and methodology, including sources and quality of the sampling frame, assessing accuracy and coverage of the sampling frame, selecting geographic units for sampling, replacing ineligible retailer sites, and addressing noncompleted inspections.

C   Calculating retailer violation rates, both weighted and unweighted.

C   Completing various required (and optional) forms, including Forms 1, 2, and 3.

C   Performing the necessary calculations to complete the forms and enter the correct data.

C   Implementing strategies to reduce the State's retailer noncompliance rate and reach the targets negotiated with the Center for Substance Abuse Prevention (CSAP).

The SPAS project staff has access to a wide range of expertise on how to implement the Synar Regulation and the Annual Synar Report. For example, many of the Single State Authorities (SSAs) have done an excellent job of responding to the Synar Regulation. SPAS can identify State staff responsible for completing this work and engage them to help other States respond to the Block Grant application and meet Synar requirements. SPAS can also engage other experts, many of whom are already listed in its expert consultant directory.

TA is available both onsite and offsite. Onsite consultants are sent directly to the State requesting the TA, and they spend the requisite amount of time consulting with the SSA staff (or the appropriate Synar-implementing agency). The SPAS project can pay for consultant travel and TA costs, as well as an honorarium, for completing the TA. For Synar TA requests, consultation is usually provided offsite, with the consultant working with the SSA staff by telephone, teleconferencing, or e-mail, whichever is most appropriate.

Although the guidelines for providing TA are fairly broad, the SPAS project does not provide consultants to write any or all of the Block Grant application. Also, there is a limit to how much TA can be provided on such issues. It is rare that a consultant is engaged for more than a day or two, and States must clearly identify their needs when making such requests.

The process for requesting TA is to contact the State's Synar reviewer at CSAP. The reviewer will assist the State in identifying specific TA needs and will then make a referral to the SPAS project staff. The SPAS staff, in consultation with the Synar reviewer, will identify the appropriate consultant and, in most cases, provide the TA within a few weeks.

Appendix C: Glossary

Area sampling. Clearly defined geographic areas constitute sampling units, which are first sampled before population elements are selected. Area sampling is used in a form of a cluster design. All of the population elements in each sampled area can be included in the survey, or a subsample of population elements can be drawn from the area.

Base sampling weight. A numerical value that is the inverse of the selection probability of a sampling unit.

Bias. Generally refers to systematic error that distorts the survey results, as opposed to random error, which balances out on average. Specifically, an estimate is said to be biased if the expected value differs from the true value.

Cluster sample. The sampling unit is a cluster; that is, a group of population elements. The more heterogeneous the population elements within clusters are, the more efficient the cluster sample becomes. In geographic area clusters, population elements tend to be homogeneous and clustering decreases the precision of survey estimates.

Completion rate. The ratio of the number of inspected outlets to the number of eligible outlets in the original sample.

Complex sample design. A sampling design that is more complex than simple random sampling. A complex design commonly incorporates disproportionate sampling, clustering, or stratification.

Confidence interval. A form of an estimator for an unknown population parameter (e.g., true violation rate) based on a random sample. It is defined as an interval (range of values, as opposed to a point estimator, which is a single number) that is accompanied by a confidence level (e.g., 95 percent). This level indicates how likely the interval includes the true unknown parameter.

Coverage. Refers to how well a sampling frame contains (covers) the target population of a survey. Two types of coverage are often mentioned, overcoverage and undercoverage. Overcoverage occurs when the frame contains elements that do not belong to the target population. Undercoverage occurs when the frame does not include all elements in the target population.

Design effect of an estimator. The ratio of the sampling variance of the estimator (e.g., violation rate estimate for the Synar survey) from a complex design to the sampling variance of the estimator under a simple random sample of the same size.

Explicit stratification. See *stratification*.

Field. The arena in which survey data are collected. The field period is the timeframe within which the field operation is carried out. Field work refers to all work conducted during the field period.

Final sample size. The number of sampled elements that are classified as survey respondents (e.g., completed inspections of eligible outlets).

Heterogeneity. The tendency that units in a group differ from one another with respect to some characteristic.

Homogeneity. The tendency that units in a group are similar to each other with respect to some characteristic.

Implicit stratification. See *stratification*

Ineligible outlet for Synar survey. An outlet that is no longer in business, does not sell tobacco, or is inaccessible to youth.

Multistage sample. A sample that is selected in more than one stage and in a nested fashion. The units selected at the first stage are called primary sampling units, units selected at the second stage are called secondary sampling units, and so forth.

Noncompleted inspection. An eligible outlet sampled for inspection is not inspected for various reasons, such as the outlet being closed at the time the inspection was attempted or possible danger to youth inspectors.

Noncompletion rate. The fraction (or percentage) of sampled eligible outlets that were not inspected. It is the ratio of the number of outlets with noncompleted inspections to the number of eligible outlets in the original sample.

Noncompliance (or violation) rate. The fraction (or percentage) of tobacco-selling outlets in a State that are accessible to minors and sell tobacco to them. The objective of the Synar survey is to estimate this rate using sampling techniques and survey inspection. It is the same as the *retailer violation rate*.

Nonsampling error. Survey error that is not attributable to sampling and can occur in every phase of survey operation.

Original sample size. The number of outlets the State selects originally for the Synar survey.

Overcoverage. Inclusion of ineligible elements in the sample frame.

Parameter. A population value that characterizes the population based on the measurements of all elements in the population. Parameters are seldom known and are estimated by statistics based on a sample drawn from the population.

Population/universe (target vs. survey). The total set of all elements being studied. The target population is the population that the survey aims to study. The survey population is the

population that is actually covered by the survey, which usually differs from the target population because the sampling frame used by the survey is imperfect.

Poststratification. Stratification created after sample selection often based on sample data. It uses the known population distribution from sources external to the survey to enhance the efficiency of a survey estimate or to benchmark the sample distribution to the known population distribution.

Primary sampling unit. Sampling unit (the highest level cluster) selected in the first stage of a multistage sample design.

Probability proportional to size sampling. A sampling method that selects a sampling unit (typically a cluster) with a probability proportional to its size. In the Synar survey, the size of the unit may be defined by the total number of outlets, the total population of youth, or some other measure appropriate to the data available on the frame.

Probability sample. A sample of units that is selected by a chance mechanism, which enables the sampler to calculate the probability of selecting the sample. To get an unbiased estimate, each sampling unit must be given nonzero probability. Compared with other nonprobability sampling methods, such as quota sampling or purposive sampling, probability sampling is the only sampling method for which the precision of the survey estimates can be measured.

Purposive sampling. A sampling method whereby population elements are selected on the basis of a subjective evaluation of how well they represent the population or some other consideration specific to the study being conducted.

Random sample. A sample selected by a chance mechanism. A synonym for probability sample.

Reserve sample. A sample drawn in addition to the original sample to be used when the target sample size is not reached because of a poor sample yield. When used, the reserve sample should be treated as if it were part of the original sample, and thus it increases the sample size. This strategy is opposed to the replacement or substitution strategy, where the sample size does not change and which can cause some bias.

Retailer violation rate. See *noncompliance rate*.

Sample. A subset of a population, as opposed to a complete census.

Sample size. The number of sampling units in a sample.

Sampling fraction. The proportion of a population that is sampled.

Sampling frame. The list of population elements that has some identification information or some device by which a sample of population elements can be selected.

Sampling unit. The unit of sample selection. In a multistage sample design, the sampling unit differs in different stages of sampling.

Sampling variance. The variance of the sampling distribution of an estimator under repeated sampling.

Secondary sampling unit. Sampling unit selected in the second stage of a multistage probability design.

Self-representing unit. A sampling unit selected with certainty. It represents only itself and is given a sampling weight of 1.

Simple random sample (SRS). A sample of a fixed size (n) is randomly selected in such a way that each possible sample of size n is equally likely to be selected. If the population size is N, there are $_N C_n$ possible samples, and each has the equal probability of selection. SRS is an equal probability sampling method, which is the simplest and most basic.

Standard deviation. The positive square root of the variance of a distribution.

Standard error. The positive square root of the sampling variance of the sampling distribution of an estimator.

Statistic. A value obtained by combining sample observations in a mathematical form. The most commonly used statistics are weighted estimates of the mean; median, proportion, and total; and corresponding variances of the estimates. Statistics can generally be divided into two types, descriptive and inferential. Descriptive statistics, such as point estimates for means and proportions, summarize sample data. Inferential statistics allow making generalizations from the sample to the population. A violation rate estimate is a descriptive statistic, whereas a confidence interval of the violation rate is an inferential statistic.

Strata. Mutually exclusive and exhaustive subgroups of a population, formed on the basis of one or more attributes related to the measure of interest (in the case of the Synar survey, the retailer violation rate). The more homogeneous the population elements are within strata, the more efficient the stratification is. During the sample selection process, outlets are independently selected *from every stratum.*

Stratification. Act of creating strata. Sometimes the term "explicit" stratification is used to refer to the usual hard boundary stratification as opposed to "implicit" stratification, where stratification is only implied through sorting the sampling frame by the stratification attributes in systematic sampling.

Subpopulation/subgroup/subclass. Any subset of a population.

Systematic random sampling. A method of randomly selecting a sample in which units are systematically selected starting from a random start and every *I*-th unit thereafter with appropriately chosen sampling interval (*I*). This method is an equal probability sampling method but is very different from the SRS method.

Tertiary sampling unit. Sampling unit selected at the third stage of a multistage sample design.

Ultimate sampling unit. Sampling unit selected in the last stage of sampling, which is an element of the population being studied.

Undercoverage. The extent to which population units are excluded from the sampling frame.

Violation rate. See *noncompliance rate*.

Weighting. Assigning numerical values to sampled elements to account for unequal probabilities of selection, noncompletion, and the deviation of sample distributions from known population distributions.

Appendix D: Checklist for Synar Survey Sampling

Step 1

_____ Develop a sampling frame (check as many sources for the frame as apply below).

_____ List sources (tobacco license list, food and/or alcohol license lists, business lists).

_____ Use area sampling frame or list-assisted area frame.

_____ Update the frame annually, as close to the time of the Synar survey as possible.

_____ Establish and implement procedures to maintain the accuracy and coverage of the frame (e.g., telephone calls, small field surveys).

Step 2

_____ Decide whether to use stratification, clustering, or both.

_____ Choose single- or multistage selection.

_____ Decide on random sampling method (simple random sampling, systematic random sampling, probability proportional to size sampling) for each stage of sampling.

_____ Exert efforts to have equal probability of selection for all tobacco outlets accessible to youth in the State.

_____ Include vending machines in the sample if they are accessible to youth.

Step 3

_____ Determine minimum required number of completed inspections to yield the right-sided 95-percent confidence interval, whose right-side boundary does not exceed 3 percentage points from the weighted violation rate estimate, taking into account the sample design.

_____ Determine the original sample size, considering past year's eligibility and completion rates (weighted rates are more appropriate for this purpose).

_____ Select a reserve sample 10 or 20 percent over the original sample size during the selection of the original sample to counter the situation where the minimum required sample size cannot be reached due to poor sample yield.

_____ If using any part of the reserve sample, add the used reserve outlets to the original sample and treat them as such. This causes an increase of the original sample size.

_____ Obtain a completion rate of at least 90 percent.

Step 4

_____ Follow the State's approved inspection protocol.

_____ Supervise youth inspectors and adult chaperones to provide rigorous monitoring and quality control.

_____ Record actual survey process in the field.

_____ Document how random selection was effected if applicable.

_____ Record the number of outlets found ineligible and specific reasons for ineligibility.

_____ Record the number of noncompleted inspections of eligible outlets and specific reasons for noncompletion.

Step 5

_____ Analyze the results of the survey, computing retailer violation rate, standard error, and 95-percent confidence interval.

_____ Use formulas and calculations appropriate to the sampling plan.

_____ Weight the results to account for unequal probabilities of selection, differences in the eligibility and completion rates among strata or clusters if applicable, and other deviations from the intended design.

_____ If vending machines are accessible to youth, compute a separate retailer violation rate, standard error, and confidence interval for them, if feasible. These outlets are usually different from over-the-counter outlets in terms of violation rate and frame accuracy.

Step 6

_____ Describe sampling methodology, including source(s) and quality of the frame; random selection process; minimum required, original, and final sample sizes; and numbers of and specific reasons for noncompleted inspections.

_____ Address any changes in sampling methodology, including the nature of the changes and the dates on which SAMHSA approval for those changes was obtained.

_____ Report retailer violation rate, rounded to the nearest 10th of a percentage point, the standard error, and the right-sided 95-percent confidence interval.

_____ Complete Form 1, Form 2 or another table showing weighting procedures and calculations, and Form 3.

_____ Verify that the numbers in Form 3 match those reported in the narrative and in Forms 1 and 2 (or other weighting table). If they do not, explain the discrepancies.

Appendix E: Bibliography

Barnett, V. (1991). *Sample Survey Principles and Methods.* New York: Oxford University Press.

Chaudhuri, A., and Vos, J. (1989). *Theory and Methods of Survey Sampling.* Amsterdam: North-Holland.

Clark, P. I.; Natanblut, S. L.; Schmitt, C. L.; Wolters, C.; and Iachan, R. (2000). Factors associated with tobacco sales to minors: lessons learned from FDA compliance checks. *JAMA* 284(6): 729-734.

Cochran, W. (1977). *Sampling Techniques.* 3rd ed. New York: John Wiley & Sons.

Dalenius, T. (1985). *Elements of Survey Sampling.* Stockholm: Swedish Agency for Research Cooperation with Developing Countries.

David, H., ed. (1978). *Contributions to Survey Sampling and Applied Statistics.* New York: Academic Press.

Deming, W. (1950). *Some Theory of Sampling.* New York: Dover Publications.

Deming, W. (1960). *Sample Design in Business Research.* New York: John Wiley & Sons.

Desu, M., and Raghavarao, D. (1990). *Sample Size Methodology.* Boston: Academic Press.

Efron, B. (1982). *The Jacknife, the Bootstrap, and Other Resampling Plans. Regional Conference Series in Applied Mathematics, No. 38.* Philadelphia: Society for Industrial and Applied Mathematics.

Fink, A. (1995). *How To Sample in Surveys.* Vol. 6, The Survey Kit. Thousand Oaks, CA: Sage Publications.

Foreman, E. (1991). *Survey Sampling Principles.* New York: Marcel Dekker.

Frankel, M. (1983). Sampling Theory. In Rossi, P.; Wright, J.; and Anderson, A. *Handbook of Survey Research.* New York: Academic Press.

Hansen, M.; Hurwitz, W.; and Madow, W. (1953). *Sample Survey Methods and Theory.* New York: John Wiley & Sons.

Hedayat, A. S., and Sinha, B. K. (1991). *Design and Inference in Finite Population Sampling.* New York: John Wiley & Sons.

Hendricks, W. A. (1956). *The Mathematical Theory of Sampling.* New Brunswick, NJ: Scarecrow.

Henry, G. (1990). *Practical Sampling.* Vol. 21, Applied Social Research Methods. Newbury Park, CA: Sage Publications.

Hess, I. (1985). *Sampling for Social Research Surveys, 1947–1980.* Ann Arbor, MI: University of Michigan Press.

Hess, I.; Riedel, D.; and Fitzpatrick, T. (1981). *Probability Sampling of Hospitals and Patients.* Ann Arbor, MI: University of Michigan Press.

Jessen, R. (1978). *Statistical Survey Techniques.* New York: John Wiley & Sons.

Johnson, N., and Smith, H., eds. (1969). *New Developments in Survey Sampling.* New York: John Wiley & Sons.

Kalton, G. (1983). *Introduction to Survey Sampling. Sage University Paper Series on Quantitative Application in the Social Sciences, No. 35.* Beverly Hills, CA: Sage Publications.

Kendall, M., and Smith, B. (1954). *Tables of Random Sampling Numbers.* New York: Cambridge University Press.

Kish, L. (1965). *Survey Sampling.* New York: John Wiley & Sons.

Kish, L. (1987). *Statistical Design for Research.* New York: John Wiley & Sons.

Konijn, H. (1973). *Statistical Theory of Sample Survey Design and Analysis.* New York: American Elsevier Publishing Co.

Krewski, D.; Platek, R.; and Rao, J., eds. (1981). *Current Topics in Survey Sampling.* New York: Academic Press.

Lavrakas, P. J. (1993). *Telephone Survey Methods: Sampling, Selection, and Supervision.* Newbury Park, CA: Sage Publications.

Lee, E.; Forthofer, R.; and Lorimer, R. (1989). *Analyzing Complex Survey Data. Sage University Paper Series on Quantitative Application in the Social Sciences, No. 71.* Beverly Hills, CA, and London: Sage Publications.

Lemeshow, S.; Hosmer, D.; and Klar, J. (1990). *Adequacy of Sample Size in Health Studies.* New York: John Wiley & Sons.

Lemeshow, S., and Stroh, G. (1985). *Sampling Techniques for Evaluating Health Parameters in Developing Countries.* Washington, DC: National Academy Press.

Lepage, R., and Billiard, L., eds. (1992). *Exploring the Limits of Bootstrap.* New York: John Wiley & Sons.

Levy, P., and Lemeshow, S. (1991). *Sampling of Populations: Methods and Applications.* New York: John Wiley & Sons.

Monroe, J., and Finkner, A. (1959). *Handbook of Area Sampling.* Philadelphia: Chilton Book Division.

Moser, C., and Kalton, G. (1971). *Survey Methods in Social Investigation.* 2d ed. New York: Basic Books.

Murthy, M. N. (1967). *Sampling Theory and Methods.* Calcutta, India: Statistical Publishing Society.

Murthy, M., et al. (1962). *Bibliography on Sampling Theory and Methods.* Calcutta, India: Indian Statistical Institute.

Namboodiri, K., ed. (1978). *Survey Sampling and Measurement.* New York: Academic Press.

Odeh, R., and Fox, M. (1991). *Sample Size Choice: Charts for Experiments With Linear Models.* 2d ed. New York: Marcel Dekker.

Raj, D. (1968). *Sampling Theory.* New York: McGraw-Hill.

Raj, D. (1972). *The Design of Sample Surveys.* New York: McGraw-Hill.

Rand Corporation (1955). *A Million Random Digits With 100,000 Normal Deviates.* Glencoe, IL: Free Press.

Rosander, A. C. (1977). *Case Studies in Sample Design.* New York: Marcel Dekker.

Salant, P., and Dillman, D. (1994). *How To Conduct Your Own Survey.* New York: John Wiley & Sons.

Sarndal, C.; Swensson, B.; and Wretman, J. (1992). *Model Assisted Survey Sampling.* New York: Springer-Verlag.

Satin, A., and Shastry, W. (1993). *Survey Sampling: A Non-Mathematical Guide.* 2d ed. Ottawa, ON: Statistics Canada.

Scheaffer, R.; Mendenhall, W.; and Ott, L. (1990). *Elementary Survey Sampling.* 4th ed. Boston: PWS-Kent Publishing Co.

Singh, D., and Chaudhary, F. (1994). *Theory and Analysis of Survey Sampling Designs.* New Delhi, India: Wiley & Sons Eastern.

Slonim, M. J. (1966). *Sampling in a Nutshell.* New York: Simon and Schuster.

Som, R. K. (1996). *Practical Sampling Techniques.* 2d ed. New York: Marcel Dekker.

Stephan, F., and McCarthy, P. (1958). *Sampling Opinions: An Analysis of Survey Procedures.* New York: John Wiley & Sons.

Stuart, A. (1976). *Basic Ideas of Scientific Sampling.* 2d ed. *Griffin's Statistical Monographs, No. 4.* London: Griffin.

Stuart, A. (1984). *The Ideas of Sampling.* High Wycombe, England: Griffin.

Sudman, S. (1976). *Applied Sampling.* New York: Academic Press.

Sukhatme, P. (1984). *Sampling Theory of Surveys with Applications.* 3rd ed. Ames, IA: Iowa State University Press.

Thompson, S. K. (1992). *Sampling.* New York: John Wiley & Sons.

U. S. Department of Health and Human Services (1994). *Preventing Tobacco Use Among Young People: A Report of the Surgeon General*. Atlanta, GA: U. S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

U. S. Department of Health and Human Services (2000). *Reducing Tobacco Use: A Report of the Surgeon General*. Atlanta, GA: U. S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health. Reprinted, with corrections, October, 2000.

Warwick, D., and Lininger, C. (1975). *The Sample Survey: Theory and Practice.* New York: McGraw-Hill.

Williams, B. (1978). *A Sampler on Sampling.* New York: John Wiley & Sons.

Wolter, K. M. (1985). *Introduction to Variance Estimation.* New York: Springer-Verlag.

Yamane, T. (1967). *Elementary Sampling Theory.* Englewood Cliffs, NJ: Prentice-Hall.

Yates, F. (1981). *Sampling Methods for Censuses and Surveys.* 4th ed. New York: Macmillan.